

# MODULAR EMBEDDINGS AND AUTOMORPHIC HIGGS BUNDLES

ROBERT A. KUCHARCZYK

**ABSTRACT.** We investigate modular embeddings for semi-arithmetic Fuchsian groups. First we prove some purely algebro-geometric or even topological criteria for a regular map from a smooth complex curve to a quaternionic Shimura variety to be covered by a modular embedding. Then we set up an adelic formalism for modular embeddings and apply our criteria to study the effect of abstract field automorphisms of  $\mathbb{C}$  on modular embeddings. Finally we derive that the absolute Galois group of  $\mathbb{Q}$  operates on the dessins d'enfants defined by principal congruence subgroups of (arithmetic or non-arithmetic) triangle groups by permuting the defining ideals in the tautological way.

## CONTENTS

1. Introduction	2
2. Uniformisation and Higgs fields	6
2.1. The upper half plane and variations of Hodge structure	6
2.2. Higgs bundles	10
2.3. Theta characteristics	11
3. Quaternionic Shimura varieties	14
3.1. The Shimura varieties themselves	14
3.2. Automorphic line bundles	17
4. Geometry of modular embeddings	20
4.1. Modular embeddings	20
4.2. Maps from curves to Shimura varieties	23
4.3. Adelic theory	26
5. Galois conjugation	27
5.1. The Serre, Taniyama and Galois groups	27
5.2. Torsors and Galois cohomology	30
5.3. Galois conjugates of Shimura varieties	30
5.4. Galois conjugates of modular embeddings	35
5.5. Application to triangle groups and dessins d'enfants	38
References	39

---

2010 *Mathematics Subject Classification.* Primary: 11F41, 11F99, 11G18, 14G35. Secondary: 11F03, 11G30, 11G35, 14G30, 14H60, 34M45.

*Key words and phrases.* Modular embeddings, Fuchsian groups, quaternionic Shimura varieties, Hilbert–Blumenthal varieties, Higgs bundles, automorphic vector bundles, variations of Hodge structure, triangle groups, dessins d'enfants, Galois conjugation.

Research partially supported by the European Research Council.

## 1. INTRODUCTION

In this article we examine several geometric and arithmetic aspects of modular embeddings in the sense of Schmutz Schaller and Wolfart [42]. These are certain holomorphic maps  $\tilde{f}: \mathbb{H} \rightarrow \mathbb{H}^r$ , where  $\mathbb{H}$  is the complex upper half plane, which are equivariant for group homomorphisms  $\varphi: \Delta \rightarrow \Gamma$ , where  $\Delta \subset \mathrm{PGL}_2(\mathbb{R})^+$  is a (not necessarily arithmetic) lattice and  $\Gamma \subset (\mathrm{PGL}_2(\mathbb{R})^+)^r$  is an irreducible arithmetic lattice, and where the first projection  $\mathrm{PGL}_2(\mathbb{R})^r \rightarrow \mathrm{PGL}_2(\mathbb{R})$  composed with  $\varphi$  induces the identity on  $\Delta$ . Such embeddings induce regular maps between algebraic varieties  $f: C \rightarrow S$ , where  $C$  is the algebraic curve with  $C^{\mathrm{an}} = \Delta \backslash \mathbb{H}$ , and  $S$  is the Shimura variety with  $S^{\mathrm{an}} = \Gamma \backslash \mathbb{H}^r$ .

As of today three series of examples are known for groups  $\Delta$  admitting modular embeddings:

- (i) If  $\Delta$  is *arithmetic* the identity  $\mathbb{H} \rightarrow \mathbb{H}$  is a modular embedding. More generally modular embeddings into higher-dimensional quaternionic Shimura varieties can be constructed quite easily.
- (ii) All Fuchsian *triangle groups*  $\Delta = \Delta_{p,q,r}$  admit modular embeddings, but only finitely many of them are arithmetic (see [51]). These modular embeddings were constructed in [5] where also modular embeddings for non-arithmetic groups were first introduced.
- (iii) Those *Veech groups* that are lattices – in which case they can also be characterised as the uniformising groups of *Teichmüller curves* – admit modular embeddings. This was discovered in [31] for genus two and in [38] for the general case.

See Examples 4.3 below for a more detailed review.

The main results of this article can be roughly divided into two distinct but related groups. In the first group we start with an arbitrary regular map  $f: C \rightarrow S$  where  $C$  is an algebraic curve over  $\mathbb{C}$  and  $S$  is a Shimura variety of the kind that appears for modular embeddings (a quaternionic connected Shimura variety, see Section 3.1 below), and give criteria of a geometric nature determining whether  $f$  arises from a modular embedding. In the second group we study the behaviour of modular embeddings under abstract field automorphisms of  $\mathbb{C}$ , and deduce a result about the Galois operation on a certain class of dessins d'enfants.

We now announce our main results in more explicit phrasing.

*Geometry of modular embeddings.* We recall the construction of irreducible arithmetic lattices in  $(\mathrm{PGL}_2(\mathbb{R})^+)^r$ . This begins with a totally real number field  $F$  and a quaternion algebra  $B/F$ . Let  $g = [F : \mathbb{Q}]$  and let  $\varrho_1, \dots, \varrho_g: F \rightarrow \mathbb{R}$  be the distinct field embeddings, and assume they are numbered in such a way that  $B$  splits at  $\varrho_1, \dots, \varrho_r$  and ramifies at  $\varrho_{r+1}, \dots, \varrho_g$ . We assume that  $r \geq 1$ , and for every  $1 \leq j \leq r$  we fix an isomorphism  $B \otimes_{F, \varrho_j} \mathbb{R} \cong \mathrm{M}_2(\mathbb{R})$ .

Let further  $\mathcal{O} \subset B$  be an order and let  $\mathcal{O}^1$  be the group of invertible elements of reduced norm one in  $\mathcal{O}$ . By the isomorphisms we just chose this embeds as an irreducible arithmetic lattice in  $\mathrm{SL}_2(\mathbb{R})^r$ , and any subgroup  $\Gamma \subset (\mathrm{PGL}_2(\mathbb{R})^+)^r$  commensurable to the image of  $\mathcal{O}^1$  in  $(\mathrm{PGL}_2(\mathbb{R})^+)^r$  is also an irreducible arithmetic lattice.

**Definition 1.1.** Let  $\Gamma$  as above, let  $\Delta \subset \mathrm{PGL}_2(\mathbb{R})^+$  be a lattice and let  $1 \leq j \leq r$ . A modular embedding for  $\Delta$  with respect to  $\varrho_j$  is a pair  $(\tilde{f}, \varphi)$  such that:

(i)  $\varphi$  is a group homomorphism  $\Delta \rightarrow \Gamma$  such that the composition

$$\Delta \xrightarrow{\varphi} \Gamma \subset (\mathrm{PGL}_2(\mathbb{R})^+)^r \xrightarrow{\mathrm{pr}_j} \mathrm{PGL}_2(\mathbb{R})^+$$

is equal to the identity inclusion (in particular,  $\Delta$  is commensurable to a subgroup of the possibly indiscrete group  $\varrho_j(\mathcal{O}^1) \subset \mathrm{PGL}_2(\mathbb{R})^+$ );

(ii)  $\tilde{f}: \mathbb{H} \rightarrow \mathbb{H}^r$  is a holomorphic map equivariant for  $\varphi: \Delta \rightarrow \Gamma$ , i.e. satisfying

$$\tilde{f}(\delta z) = \varphi(\delta) \tilde{f}(z) \quad \text{for all } z \in \mathbb{H}, \delta \in \Delta.$$

As mentioned above, such a modular embedding defines a regular map of algebraic varieties  $f: C \rightarrow S$ . Our first two main results give purely algebro-geometric (or even topological) criteria when a morphism from a curve to a quaternionic Shimura variety is covered by a modular embedding. We need to distinguish between two cases, depending on whether  $C$  is affine or projective.

To state the results it is necessary to assume that  $\Gamma$  is torsion-free; this can always be assumed by passing to a finite index subgroup, which is benign for our purposes. The cotangent bundle  $\Omega_{\mathbb{H}^r}^1$  splits in an obvious way as a sum of line bundles, induced by the product decomposition  $\mathbb{H}^r = \prod_{j=1}^r \mathbb{H}$ . Since the elements of  $\Gamma$  preserve this decomposition, it descends to the cotangent bundle of  $S$ :

$$\Omega_S^1 = \bigoplus_{j=1}^r \mathcal{M}_j$$

for some algebraic line bundles  $\mathcal{M}_j$  on  $S$ .

**Theorem A1.** Let  $S$  be a Shimura variety of the kind just described, constructed from a quaternion algebra  $B/F$ , let  $C$  be a smooth projective complex curve, let  $f: C \rightarrow S$  be a morphism of complex algebraic varieties and let  $\varrho_j$  be an archimedean place of  $F$  at which  $B$  is unramified. Then the following are equivalent:

- (i)  $f$  is covered by a modular embedding with respect to  $\varrho_j$ .
- (ii)  $f^* \mathcal{M}_j$  is isomorphic to the canonical bundle  $\omega_C$  as an algebraic line bundle.
- (iii)  $f^* \mathcal{M}_j$  is isomorphic to  $\omega_C$  as a topological line bundle (i.e. the degree of  $f^* \mathcal{M}_j$  is minus the Euler characteristic of  $C$ ).

This is Corollary 4.5 below. An extension of Theorem A1, with some additional equivalent conditions, is given and proved below as Theorem 4.4.

Theorem A1 is proved using simple special cases of ideas from C. Simpson's correspondence between Higgs bundles and local systems. The line bundles  $\mathcal{M}_j$  come with distinguished square roots  $\mathcal{L}_j$  that can be viewed as bundles of certain modular forms for  $\Gamma$ , see Remark 3.3 below. Phrased more abstractly, they are automorphic line bundles. For a map  $f$  covered by a modular embedding,  $f^* \mathcal{L}_j$  is then a theta characteristic, hence there is a natural isomorphism  $\vartheta_j: f^* \mathcal{L}_j \rightarrow f^* \mathcal{L}_j^{-1} \otimes \omega_C^1$ , i.e. a *maximal Higgs field*. This field can also be constructed without the assumption that  $f$  is covered by a modular embedding, and its maximality is equivalent to  $f$  being covered by a modular embedding.

In the affine case we obtain a very similar result. Then the line bundles  $f^* \mathcal{L}_j$  can be constructed in a different way. If  $C$  is affine and  $f: C \rightarrow S$  is covered by

a modular embedding then  $S$  is automatically a *Hilbert–Blumenthal variety*, i.e. a quaternionic Shimura variety derived from the quaternion algebra  $B = M_2(F)$  (see Proposition 4.6 below). These varieties are moduli spaces for abelian varieties with real multiplication by  $F$ , hence a morphism  $f: C \rightarrow S$  (modular embedding or not) corresponds to a family  $p_C: A_C \rightarrow C$  of such abelian varieties. To such a family we attach its *Hodge bundle*  $p_{C,*}\Omega_{A_C/C}^1$ , a vector bundle on  $C$  with a  $\mathbb{Q}$ -linear action by  $F$ . Fixing an archimedean place  $\varrho_j$  of  $F$  we denote the  $\varrho_j$ -eigenspace in the Hodge bundle by  $\mathcal{E}$ ; this is an algebraic line bundle on  $C$ .

Letting  $\overline{C}$  be a projective completion of  $C$  with boundary  $D$  the theory of *variations of Hodge structure* provides us with a canonical extension  $\mathcal{E}_{\text{ext}}$  of  $\mathcal{E}$  to  $\overline{C}$ . Our main geometric result in this case is then the following.

**Theorem A2.** *Let  $C = \overline{C} \setminus D$  be a smooth affine curve, let  $p_C: A_C \rightarrow C$  be a family of abelian varieties with real multiplication by  $F$ , let  $f: C \rightarrow S$  be the corresponding regular map where  $S$  is a Hilbert–Blumenthal variety, and let  $\varrho_j: F \rightarrow \mathbb{R}$  be an archimedean place. Let  $\eta_{\text{ext}}$  and  $\mathcal{E}_{\text{ext}}$  be as above. Then the following are equivalent:*

- (i)  *$f$  is covered by a modular embedding for  $\varrho$ .*
- (ii)  *$\mathcal{E}_{\text{ext}}$  is a logarithmic theta characteristic on  $C$ , i.e. its square is isomorphic to  $\omega_{\overline{C}}(\log D)$ .*
- (iii) *The degree of  $\mathcal{E}_{\text{ext}}$  is  $-\frac{1}{2}$  times the Euler characteristic of  $C$ .*

This is proved as Theorem 4.8 below. Note the similarity to Theorem A1: the restriction of  $\mathcal{E}_{\text{ext}}$  to  $C$  is  $f^*\mathcal{L}_j$ , a square root of  $f^*\mathcal{M}_j$ . Compare also Remark 4.9 below.

There are other results that express a seemingly transcendental condition (as being covered by a modular embedding presents itself) on a regular map between algebraic varieties in a purely algebro-geometric language and that also proceed via Simpson’s correspondence. To the present author’s knowledge, this idea was first used in [53] to study Shimura curves in the moduli space  $\mathcal{A}_g$  of principally polarised abelian varieties, then for Teichmüller curves in [38] and for Kobayashi geodesics in  $\mathcal{A}_g$  in [39].

*Field automorphisms.* A consequence of Theorems A1 and A2 is that the property of being covered by a modular embedding is stable under field automorphisms of  $\mathbb{C}$ . For  $\tau \in \text{Aut } \mathbb{C}$  the variety  $\tau S$  is again a quaternionic Shimura variety and the bundle  $\tau\mathcal{M}_j$  occurs again as a summand of the cotangent bundle. This implies:

**Theorem B.** *Let  $C$  and  $S$  be as in Theorem A1 or A2, let  $f: C \rightarrow S$  be covered by a modular embedding with respect to  $\varrho_j: F \rightarrow \mathbb{R}$  and let  $\tau \in \text{Aut } \mathbb{C}$ . Then  $\tau f: \tau C \rightarrow \tau S$  is covered by a modular embedding with respect to  $\tau \circ \varrho_j$ .*

This is Theorem 5.9 below. A more refined version, employing the adelic formalism of Shimura varieties and thus giving a more explicit version of this result, is formulated as Theorem 5.12. It makes use of the explicit theory of Galois conjugates of Shimura varieties in the form conjectured by Langlands [28] and proved by Milne and Shih [32, 36], and its extension by Milne to automorphic vector bundles [33].

*Dessins d'enfants.* Theorem B (or a more precise form thereof) has applications to the theory of dessins d'enfants. Let  $p, q, r$  be positive integers or  $\infty$  such that

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} < 1$$

(where we set  $\frac{1}{\infty} = 0$ ), and let  $\Delta_{p,q,r} \subset \mathrm{PGL}_2(\mathbb{R})^+$  be the triangle group of signature  $(p, q, r)$ . This group admits a modular embedding to an arithmetic group defined by a certain quaternion algebra over the field

$$K_{p,q,r} = \mathbb{Q}(\mathrm{tr} \Delta_{p,q,r}) = \mathbb{Q}\left(\cos \frac{\pi}{p}, \cos \frac{\pi}{q}, \cos \frac{\pi}{r}\right),$$

and we may define principal congruence subgroups  $\Delta_{p,q,r}(\mathfrak{a})$  for every integral ideal  $\mathfrak{a}$  of  $K_{p,q,r}$ . The natural projection

$$\Delta_{p,q,r}(\mathfrak{a}) \backslash \mathbb{H} \rightarrow \Delta_{p,q,r} \backslash \mathbb{H} \cong \mathbb{P}^1(\mathbb{C})$$

is then unramified outside  $\{0, 1, \infty\} \subset \mathbb{P}^1(\mathbb{C})$ , i.e. it is a *Belyi map* and therefore corresponds to a dessin  $\mathcal{D}_{p,q,r}(\mathfrak{a})$ .

**Theorem C.** *The action by  $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$  on  $\mathcal{D}_{p,q,r}(\mathfrak{a})$  is given by*

$$\tau \mathcal{D}_{p,q,r}(\mathfrak{a}) = \mathcal{D}_{p,q,r}(\tau \mathfrak{a}).$$

This is Theorem 5.14 below. Several special cases have been proved before by rather different methods, cf. the discussion following Theorem 5.14.

*Outline.* In Sections 2 and 3 we summarise known or essentially known material that is used in proving our main theorems but rather scattered throughout the literature, so we recall it here in phrasings suitable for our purposes. Section 2 contains a discussion of variations of Hodge structure, Higgs fields and Simpson's correspondence in the very special one-dimensional cases that we need, and some results about their relation to geometry, most importantly Proposition 2.8, the main nontrivial result used to prove Theorems A1 and A2. In Section 3 we recall the formalism of connected Shimura varieties and automorphic line bundles, with special emphasis on quaternionic Shimura varieties, which are those that can occur as targets for modular embeddings.

In Section 4 we introduce modular embeddings, discuss examples and recall some of their properties, leading to proofs of Theorems A1 and A2. The effect of field automorphisms of  $\mathbb{C}$  is discussed in Section 5, beginning with a summary of Deligne–Langlands–Milne–Shih's results on conjugates of Shimura varieties and automorphic bundles and continuing with applications to modular embeddings, including Theorems B and C.

*Acknowledgements.* This article started as a chapter in [27]. The author wishes to thank his coauthor, John Voight, for suggesting to publish it as a separate article, and Jürgen Wolfart for encouraging him to try to prove Theorem C using modular embeddings. Furthermore he wishes to thank these two people and in addition Ursula Hamenstädt, Martin Möller, Joachim Schwermer, Jakob Stix, Martin Westerholt-Raum and Jonathan Zachhuber for insightful conversations about the contents of this article and surrounding topics. Last but not least thanks go to Allan Hale for printing.

*Notation.* For a field  $k$  the symbol  $\bar{k}$  denotes an algebraic closure of  $k$ . Specifically,  $\overline{\mathbb{Q}}$  denotes *the* algebraic closure of  $\mathbb{Q}$  in  $\mathbb{C}$ . Complex conjugation, or its restriction to any subfield of  $\mathbb{C}$ , is denoted by  $\iota$ . The profinite completion of  $\mathbb{Z}$  is denoted by  $\hat{\mathbb{Z}} = \prod_{p \text{ prime}} \mathbb{Z}_p$  and the ring of finite rational adèles is denoted by  $\mathbb{A}^f = \hat{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbb{Q}$ .

For a field extension  $E/F$ , Weil's restriction of scalars is denoted by  $\text{Res}_{E/F}$ . The multiplicative group over a field  $F$  is denoted by  $\mathbb{G}_m = \mathbb{G}_{m,F} = \text{GL}_{1,F}$ . We often write  $\mathbb{T}^E$  for  $\text{Res}_{E/F} \mathbb{G}_m$  when  $F$  is clear from the context. For a linear algebraic group  $G$  over a field  $k$  we write  $X^*(G) = \text{Hom}(G_{\bar{k}}, \mathbb{G}_{m,\bar{k}})$  and  $X_*(G) = \text{Hom}(\mathbb{G}_{m,\bar{k}}, G_{\bar{k}})$ , both regarded as  $\text{Gal}(\bar{k}/k)$ -modules.

For an algebraic group  $G$  defined over a subfield  $k$  of  $\mathbb{R}$  we denote the connected component of the identity in  $G(\mathbb{R})$  by  $G(\mathbb{R})^+$ , and we set  $G(k)^+ = G(k) \cap G(\mathbb{R})^+$ .

The group of  $n$ -th roots of unity (as an abstract group scheme, or as a subgroup of  $\mathbb{C}$ ) is denoted by  $\mu_n$ .

Holomorphic or algebraic vector bundles on a complex manifold or variety  $X$  are denoted by latin letters such as  $V$  when considered as fibre bundles  $V \rightarrow X$  with fibrewise vector space structure, and by corresponding calligraphic letters such as  $\mathcal{V}$  when considered as sheaves of  $\mathcal{O}_X$ -modules.

If  $X$  is a variety over  $\mathbb{C}$ , the associated complex analytic space is denoted by  $X^{\text{an}}$ .

Two different symbols for isomorphism are used:  $X \simeq Y$  denotes the existence of an isomorphism  $X \rightarrow Y$  and  $X \cong Y$  denotes that a particular (usually canonical) isomorphism has been chosen.

If  $f: X \rightarrow Y$  is a map and  $A \subseteq X$  and  $B \subseteq Y$  are subsets with  $f(A) \subseteq B$ , then  $f|_A^B$  denotes the induced map  $A \rightarrow B$ , and  $f|A$  means  $f|_A^A$  when it makes sense.

For two sets  $A$  and  $B$  the symmetric difference is denoted by  $A \triangle B = (A \cup B) \setminus (A \cap B)$ .

For an equivalence relation on a set  $X$  the equivalence class containing  $x \in X$  is denoted by  $[x]$ ; for equivalence classes in products  $X \times Y$  we write  $[x, y]$  instead of  $[(x, y)]$ .

## 2. UNIFORMISATION AND HIGGS FIELDS

By the uniformisation theorem the universal cover  $\tilde{X}$  of every finite type Riemann surface  $X$  with negative Euler characteristic is biholomorphic to the upper half plane  $\mathbb{H} = \{z \in \mathbb{C} \mid \text{Im } z > 0\}$ , hence  $X$  is biholomorphic to a quotient  $\Delta \backslash \mathbb{H}$  where  $\Delta$  is a torsion-free lattice in the biholomorphism group  $\text{Aut } \mathbb{H}$ . We identify this group with  $\text{PGL}_2(\mathbb{R})^+$  acting via Möbius transformations.<sup>(\*)</sup> The geometry behind this result is much subtler than it appears at first sight, see e.g. [17, 22].

The upper half plane naturally occurs in many different contexts. For the purpose of the present work it is convenient to interpret it as a *period domain for Hodge structures*.

**2.1. The upper half plane and variations of Hodge structure.** We recall the basic definitions around Hodge structures, cf. [8, 21, 41]. Let  $V$  be a finite-dimensional real vector space. A *real Hodge structure* on  $V$  is a decomposition of

---

<sup>(\*)</sup>This is more commonly called  $\text{PSL}_2(\mathbb{R})$ , but we wish to avoid this notation since it wrongly suggests the existence of a linear algebraic group  $\text{PSL}_2$  whose group of real-valued points is  $\text{PGL}_2(\mathbb{R})^+$ .

$V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$  into complex vector subspaces,

$$V_{\mathbb{C}} = \bigoplus_{p,q \in \mathbb{Z}} V^{p,q},$$

such that  $V^{p,q}$  is the complex conjugate of  $V^{q,p}$ . The *Hodge numbers* of this structure are the numbers  $v^{p,q} = \dim_{\mathbb{C}} V^{p,q}$ . There are two important data derived from this decomposition:

- (i) the *weight decomposition*  $V = \bigoplus_{n \in \mathbb{Z}} V_n$  (direct sum of real vector spaces), where  $V_{n,\mathbb{C}} = \bigoplus_{p+q=n} V^{p,q}$ , and
- (ii) the *Hodge filtration*  $F^{\bullet} V_{\mathbb{C}}: \cdots \supseteq F^{-1} V_{\mathbb{C}} \supseteq F^0 V_{\mathbb{C}} \supseteq F^1 V_{\mathbb{C}} \supseteq \cdots$  defined by  $F^p V_{\mathbb{C}} = \bigoplus_{p' \geq p} V^{p',q}$ .

The weight decomposition and the Hodge filtration together suffice to reconstruct the Hodge structure as

$$V^{p,q} = V_{\mathbb{C}, p+q} \cap F^p V_{\mathbb{C}} \cap \overline{F^q V_{\mathbb{C}}}.$$

A Hodge structure is called *pure of weight  $n$*  if  $V = V_n$ . The standard example of a Hodge structure of pure weight  $n$  is  $H^n(X, \mathbb{R})$  for a compact Kähler manifold  $X$ .

These definitions can be rephrased in terms of representations of the *Serre torus*  $\mathbb{S} = \text{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_m$  with  $\mathbb{S}(\mathbb{R}) = \mathbb{C}^{\times}$ . If  $V$  is a finite-dimensional real vector space and  $\mathbb{S} \rightarrow \text{GL}(V)$  is a rational representation, we let  $V^{p,q}$  be the subspace of  $V_{\mathbb{C}}$  where  $z \in \mathbb{C}^{\times} = \mathbb{S}(\mathbb{R})$  operates as multiplication by  $z^p \bar{z}^q$ . This assignment defines an equivalence of categories between representations of  $\mathbb{S}$  and real Hodge structures.

The weight decomposition corresponds to the homomorphism  $w: \mathbb{G}_{m,\mathbb{R}} \rightarrow \mathbb{S}$  given on real-valued points by  $r \mapsto r^{-1}$ . Hence the weight decomposition is characterised by the fact that  $w(r)$  operates as  $r^{-n}$  on  $V_n$ .

The Hodge filtration corresponds to the homomorphism  $\mu: \mathbb{G}_{m,\mathbb{C}} \rightarrow \mathbb{S}_{\mathbb{C}}$  given on complex-valued points by  $\mathbb{C}^{\times} \rightarrow \mathbb{C}^{\times} \times \mathbb{C}^{\times}$ ,  $z \mapsto (z, 1)$ . Here we use the unique isomorphism of complex Lie groups  $\mathbb{S}(\mathbb{C}) \rightarrow \mathbb{C}^{\times} \times \mathbb{C}^{\times}$  such that the composition

$$\mathbb{C}^{\times} = \mathbb{S}(\mathbb{R}) \hookrightarrow \mathbb{S}(\mathbb{C}) \cong \mathbb{C}^{\times} \times \mathbb{C}^{\times}$$

is the identity on the first coordinate and complex conjugation on the second one.

For this reason, whenever we have a homomorphism  $h: \mathbb{S} \rightarrow G$  where  $G$  is a real algebraic group, we think of it as a ‘Hodge structure on  $G$ ’ and call  $w_h = h \circ w: \mathbb{G}_{m,\mathbb{R}} \rightarrow G$  the ‘weight decomposition’ and  $\mu_h = h \circ \mu: \mathbb{G}_{m,\mathbb{C}} \rightarrow G$  the ‘Hodge filtration’.

Tensor products, duals, symmetric powers and alternating powers of Hodge structures are defined in the obvious ways. For each  $n \in \mathbb{Z}$  we define the *Tate Hodge structure*<sup>(†)</sup>  $\mathbb{R}(n)$  of pure weight  $-2n$  that corresponds to the representation  $\mathbb{S} \rightarrow \text{GL}(1)$  sending  $z \in \mathbb{C}^{\times} = \mathbb{S}(\mathbb{R})$  to  $|z|^{-n}$ . Its underlying real vector space is usually written as  $(2\pi i)^n \mathbb{R}$ , and for nonnegative  $n$  there is a canonical isomorphism  $\mathbb{R}(-n) \cong H^{2n}(\mathbb{P}^n(\mathbb{C}), \mathbb{R})$ .

<sup>(†)</sup>Not to be confused with Hodge–Tate structures, which are certain  $p$ -adic Galois representations.

*Polarisations.* For a real Hodge structure  $V$  with corresponding representation  $\varrho: \mathbb{S} \rightarrow \mathrm{GL}(V)$  the *Weil operator* on  $V$  is the real linear map  $C: V \rightarrow V$  that acts as  $i^{p-q}$  on  $V^{p,q}$ . It can also be described as  $\varrho(i)$  where  $i$  is interpreted as an element of  $\mathbb{C}^\times = \mathbb{S}(\mathbb{R})$ .

Let  $V$  be a real Hodge structure which is pure of weight  $n$ . A *polarisation* of  $V$  is a morphism of Hodge structures  $\psi: \bigwedge^2 V \rightarrow \mathbb{R}(-n)$  for which the symmetric pairing

$$\Psi: V_{\mathbb{R}} \times V_{\mathbb{R}} \rightarrow \mathbb{R}, \quad (v, w) \mapsto (2\pi i)^n \psi(v, Cw),$$

is positive definite. The Hodge structures  $H^n(X^{\mathrm{an}}, \mathbb{R})$  for smooth projective varieties  $X$  admit polarisations coming from ample line bundles on  $X$ .

*Variations.* If  $X$  is a complex manifold, a *variation of (real) Hodge structure*  $\mathbb{V}$  on  $X$  consists of a local system  $\mathbb{V}_{\mathbb{R}}$  of real vector spaces on  $X$  together with a Hodge structure on each fibre  $\mathbb{V}_{\mathbb{R}}(x)$  such that the weight decomposition is induced by a decomposition of local systems  $\mathbb{V}_{\mathbb{R}} = \bigoplus_{n \in \mathbb{Z}} \mathbb{V}_{\mathbb{R}, n}$ , the Hodge filtration is induced by a filtration of holomorphic vector bundles  $F^\bullet \mathcal{V}: \dots \supseteq F^{-1} \mathcal{V} \supseteq F^0 \mathcal{V} \supseteq F^1 \mathcal{V} \supseteq \dots$  and Griffiths's transversality condition  $\nabla F^p \mathcal{V} \subseteq F^{p-1} \mathcal{V} \otimes \Omega_X^1$  holds. Here  $\mathcal{V} = \mathbb{V}_{\mathbb{R}} \otimes_{\mathbb{R}} \mathcal{O}_X$  is the holomorphic vector bundle underlying  $\mathbb{V}_{\mathbb{C}}$  and  $\nabla: \mathcal{V} \rightarrow \mathcal{V} \otimes \Omega_X^1$  is the unique holomorphic integrable connection for which the sections of  $\mathbb{V}_{\mathbb{C}}$  are flat.

Morphisms, tensor products and so forth are defined in the obvious way, as is the 'constant variation'  $\underline{V}_X$  associated with a Hodge structure  $V$ . A *polarisation* of a variation  $\mathbb{V}$  of pure weight  $n$  is a morphism of variations  $\bigwedge^2 \mathbb{V} \rightarrow \underline{\mathbb{R}(-n)}_X$  which is a polarisation fibrewise. Usually the existence of a polarisation on each weight component is included in the definition of a variation of Hodge structure.

Again the motivating example is derived from the cohomology of algebraic varieties: let  $f: Y \rightarrow X$  be a smooth projective morphism of algebraic varieties, then there is a variation of real Hodge structure  $\mathbb{V}$  with underlying local system  $\mathbb{V}_{\mathbb{R}} = R^n f_* \mathbb{R}$ ; the fibre at  $x \in X$  is  $H^n(f^{-1}(x)^{\mathrm{an}}, \mathbb{R})$  equipped with its natural Hodge structure. These variations admit polarisations that are constructed from relatively ample line bundles on  $Y$ .

*The upper half plane as a period domain.* Let  $W = \mathbb{R}^2$  with the symplectic form

$$\psi: W \times W \rightarrow \mathbb{R}(-1) = \frac{1}{2\pi i} \mathbb{R}, \quad \left( \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right) \mapsto \frac{1}{2\pi i} (v_1 w_2 - v_2 w_1). \quad (1)$$

Then  $\mathbb{H}$  parameterises Hodge structures on  $W$  for which  $\psi$  is a polarisation and which have Hodge numbers  $v^{1,0} = v^{0,1} = 1$  and all other  $v^{p,q} = 0$ . An easy calculation shows that all such Hodge structures have the form

$$F^1 W_{\mathbb{C}} = W^{1,0} = \mathbb{C} \cdot \begin{pmatrix} z \\ 1 \end{pmatrix} \quad \text{and} \quad W^{0,1} = \mathbb{C} \cdot \begin{pmatrix} \bar{z} \\ 1 \end{pmatrix}$$

for a uniquely determined  $z \in \mathbb{H}$ .

These Hodge structures define a polarised variation of Hodge structure  $\mathbb{W}$  on  $\mathbb{H}$ , given by the following data:

- (i) the trivial local system of  $\mathbb{R}$ -vector spaces  $\mathbb{W}_{\mathbb{R}} = \underline{W}$  on  $\mathbb{H}$ ;
- (ii) the symplectic form  $\psi: \bigwedge^2 \mathbb{W}_{\mathbb{R}} \rightarrow \underline{\mathbb{R}(-1)}$ ;

- (iii) the filtration of the trivial holomorphic vector bundle  $\mathcal{W} = \mathbb{W}_{\mathbb{R}} \otimes_{\mathbb{R}} \mathcal{O}_{\mathbb{H}} \cong \mathcal{O}_{\mathbb{H}}^2$  with  $F^p \mathcal{W} = \mathcal{W}$  for  $p \leq 0$ ,

$$F^1 \mathcal{W} \subset \mathcal{W}, \quad (F^1 \mathcal{W})_z = \mathbb{C} \cdot \begin{pmatrix} z \\ 1 \end{pmatrix},$$

and  $F^0 \mathcal{W} = 0$  for  $p \geq 2$ .

The previous considerations show:

**Lemma 2.1.** *Let  $Y$  be a complex manifold. Then there is a natural bijection between the following two types of objects:*

- (i) *polarised variations of real Hodge structure  $\mathbb{V}$  on the trivial local system  $\mathbb{V}_{\mathbb{R}} = \underline{\mathbb{R}}^2_Y$  with the symplectic form (1) with Hodge numbers  $v^{1,0} = v^{0,1} = 1$  and all other  $v^{p,q} = 0$ ;*
- (ii) *holomorphic maps  $f: Y \rightarrow \mathbb{H}$ .*

*This bijection is effectuated as follows:  $f$  as in (ii) corresponds to the variation  $f^* \mathbb{W}$  on  $Y$ , and a variation  $\mathbb{V}$  corresponds to the period map*

$$f: Y \rightarrow \mathbb{H}, \quad y \mapsto z \text{ such that } (F^1 \mathcal{W})_y = \mathbb{C} \cdot \begin{pmatrix} z \\ 1 \end{pmatrix}.$$

The universal variation  $\mathbb{W}$  on  $\mathbb{H}$  admits an action by  $\mathrm{SL}_2(\mathbb{R})$  extending that on  $\mathbb{H}$  by Möbius transformations: if  $W = \mathbb{H} \times \mathbb{R}^2$  is the total space of  $\mathbb{W}_{\mathbb{R}}$ , we set

$$g: W \rightarrow W, \quad (z, w) \mapsto (gz, gw),$$

for  $g \in \mathrm{SL}_2(\mathbb{R})$ , where  $z \mapsto gz$  is the usual operation by Möbius transformations and  $w \mapsto gw$  is the usual matrix-by-vector multiplication. A direct computation shows that the action of  $\mathrm{SL}_2(\mathbb{R})$  on  $\mathbb{W}_{\mathbb{C}}$  by  $\mathbb{C}$ -linear extension preserves the Hodge filtrations, and therefore this action can be viewed as an action on  $\mathbb{W}$  as a polarised variation of real Hodge structure.

If  $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$  is a torsion-free discrete subgroup, then  $\mathbb{W}$  descends to a polarised variation of real Hodge structure  $\mathbb{W}_{\Gamma} = \Gamma \backslash \mathbb{W}$  on  $X_{\Gamma} = \Gamma \backslash \mathbb{H}$  which we call the *Fuchsian variation* on  $X_{\Gamma}$ .

*Remark 2.2.* If  $\Gamma \subset \mathrm{SL}_2(A)$  for some subring  $A \subseteq \mathbb{R}$ , the Fuchsian variation admits a natural  $A$ -structure  $\mathbb{W}_{\Gamma, A}$ . In particular for torsion-free subgroups  $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$  we obtain a principally polarised variation of *integral* Hodge structure on  $X_{\Gamma}$ ; it is isomorphic to the variation  $R^1 p_* \mathbb{Z}$  for the universal or ‘modular’ family of elliptic curves  $p: E_{\Gamma} \rightarrow X_{\Gamma}$ , and the family can be reconstructed from the variation. In this sense the Fuchsian variation  $\mathbb{W}_{\Gamma}$  serves as a replacement for this family of elliptic curves also for the cases when  $\Gamma$  is not contained in  $\mathrm{SL}_2(\mathbb{Z})$ .

**Definition 2.3.** *Let  $X$  be a complex manifold. A real variation of elliptic curve type on  $X$  is a polarised variation of real Hodge structure  $\mathbb{V}$  on  $X$  with Hodge numbers  $v^{1,0} = v^{0,1} = 1$  and all other  $v^{p,q} = 0$ .*

For such a variation the local system  $\mathbb{V}_{\mathbb{R}}$  need not be trivial. However, its pullback to the universal cover  $\tilde{X}$  is. After choosing a trivialisation of this pullback we obtain a period map  $\tilde{X} \rightarrow \mathbb{H}$ . This trivialisation allows us to interpret the monodromy of the local system on  $X$  as a representation  $\varrho: \pi_1(X) \rightarrow \mathrm{SL}_2(\mathbb{R})$ ; the period map

is equivariant for  $\varrho$ . If the image of  $\varrho$  is a *discrete* subgroup  $\Gamma$  of  $\mathrm{SL}_2(\mathbb{R})$ , then the period map descends to a holomorphic map  $f: X \rightarrow \Gamma \backslash \mathbb{H}$ , and  $\mathbb{V} \cong f^* \mathbb{W}_\Gamma$ .

*Deligne's canonical extension.* Let  $\overline{X}$  be a smooth compact complex manifold and let  $Y \subset \overline{X}$  be a normal crossings divisor (these conditions are met if  $\overline{X}$  is a compact Riemann surface and  $Y$  is a finite subset). Set  $X = \overline{X} \setminus Y$ , and let  $\mathbb{V}$  be a polarised variation of real Hodge structure on  $X$ . Assume furthermore that the monodromy around  $Y$  operates by unipotent endomorphisms on the fibres of  $\mathbb{V}_\mathbb{R}$ . This is true, for instance, if  $X = \Gamma \backslash \mathbb{H}$  and  $\mathbb{V} = \mathbb{W}_\Gamma$  for a torsion-free lattice  $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$  that does not contain parabolic elements with eigenvalue  $-1$ . Let  $\mathcal{V} = \mathbb{V}_\mathbb{R} \otimes_{\mathbb{C}} \mathcal{O}_X$  be the vector bundle underlying the variation, and let  $\nabla: \mathcal{V} \rightarrow \mathcal{V} \otimes \Omega_X^1$  be the natural connection.

Then there exist an extension of  $\mathcal{V}$  to a holomorphic vector bundle  $\mathcal{V}_{\mathrm{ext}}$  on  $\overline{X}$  and an extension of  $\nabla$  to a logarithmic connection, i.e. a sheaf homomorphism  $\nabla_{\mathrm{ext}}: \mathcal{V}_{\mathrm{ext}} \rightarrow \mathcal{V}_{\mathrm{ext}} \otimes \Omega_{\overline{X}}^1(\log Y)$  satisfying the Leibniz rule, such that the residue of  $\nabla_{\mathrm{ext}}$  at each component  $Y^0$  of  $Y$  is a nilpotent endomorphism of  $\mathcal{V}_{\mathrm{ext}}|_{Y^0}$ . The pair  $(\mathcal{V}_{\mathrm{ext}}, \nabla_{\mathrm{ext}})$  is determined uniquely up to isomorphism by these properties. Its existence and uniqueness are proved in [7, II.§5].

The subbundles  $F^p \mathcal{V} \subseteq \mathcal{V}$  extend uniquely to holomorphic subbundles  $F^p \mathcal{V}_{\mathrm{ext}} \subseteq \mathcal{V}_{\mathrm{ext}}$ . This is a direct consequence of the Nilpotent Orbit Theorem [41, (4.9)].

*Remark 2.4.* Assume that  $\overline{X}$  is a smooth projective variety, and that  $\mathbb{V}$  is isomorphic to the variation  $R^n f_* \mathbb{R}$  for some smooth proper morphism  $f: Y \rightarrow X$ . Then the vector bundle  $\mathcal{V}$  has an algebraic structure: it can be identified with the relative de Rham cohomology  $\mathcal{H}_{\mathrm{dR}}^1(Y/X) = R^1 f_* \Omega_{Y/X}^\bullet$ , and the subbundles  $F^p \mathcal{V}$  can be constructed by truncating the relative de Rham complex  $\Omega_{Y/X}^\bullet$ ; hence they, too, are algebraic subbundles of  $\mathcal{V}$ . But there is also a different way to find algebraic structures on these bundles: the canonical extensions  $\mathcal{V}_{\mathrm{ext}}$  and  $F^p \mathcal{V}_{\mathrm{ext}}$  are holomorphic vector bundles on  $\overline{X}$ , hence by Serre's 'GAGA' theorems [44, Théorèmes 2 et 3] they possess unique algebraic structures which we can restrict to  $X$ . By [41, (4.13)] these two constructions produce the same algebraic structures on  $\mathcal{V}$  and its Hodge subbundles.

**2.2. Higgs bundles.** It is often convenient to work with systems of Hodge bundles (see [48, p. 898]), a souped-up version of Higgs bundles, instead of variations of Hodge structure, since they fit more naturally into the realm of algebraic varieties. By deep results of C. Simpson [48, 49, 50], variations of Hodge structure and systems of Hodge bundles contain the same information in many important cases, including our setup. For an introduction to this circle of ideas see [29]. We need only an elementary special case, for which the whole machinery is not necessary.

**Definition 2.5.** (i) Let  $X$  be a complex manifold or an algebraic variety (over an arbitrary base field). A Higgs bundle of elliptic curve type on  $X$  consists of a (holomorphic or algebraic) line bundle  $\mathcal{L}$  on  $X$  and an  $\mathcal{O}_X$ -linear map  $\vartheta: \mathcal{L} \rightarrow \mathcal{L}^{-1} \otimes_{\mathcal{O}_X} \Omega_X^1$ , called the Higgs field.

(ii) If  $\overline{X}$  is a smooth compact complex manifold or a smooth complete algebraic variety and  $Y \subset \overline{X}$  is a normal crossings divisor, a logarithmic Higgs

bundle of elliptic curve type on  $X = \bar{X} \setminus Y$  is a line bundle  $\mathcal{L}$  on  $\bar{X}$  together with an  $\mathcal{O}_{\bar{X}}$ -linear map  $\vartheta: \mathcal{L} \rightarrow \mathcal{L}^{-1} \otimes \Omega_{\bar{X}}^1(\log Y)$ .

Let  $X$  be a complex manifold and let  $\mathbb{V}$  be a real variation of elliptic curve type on  $X$ . There is a simple way to construct a Higgs bundle from  $\mathbb{V}$ : Set  $\mathcal{L} = F^1\mathcal{V}$  and (for the moment)  $\mathcal{M} = \mathcal{V}/\mathcal{L}$ . The fibre of  $\mathcal{M}$  over  $x \in X$  is  $\mathbb{V}_{x,\mathbb{C}}/\mathbb{V}_x^{1,0}$  which is canonically isomorphic to  $\mathbb{V}_x^{0,1}$ ; note, however, that the  $\mathbb{V}_x^{0,1}$  for varying  $x$  form not a holomorphic but an antiholomorphic subbundle of  $\mathbb{V}_{\mathbb{C}}$ , which is isomorphic as a  $\mathcal{C}^\infty$ -vector bundle to  $\mathcal{M} \otimes_{\mathcal{O}_X} \mathcal{C}_Y^\infty$ .

**Lemma 2.6.** *With this notation,  $\mathcal{M}$  is canonically isomorphic to the dual of  $\mathcal{L}$ .*

Hence we may identify  $\mathcal{M}$  with  $\mathcal{L}^{-1}$ .

*Proof.* The short exact sequence

$$0 \rightarrow \mathcal{L} \rightarrow \mathcal{V} \rightarrow \mathcal{M} \rightarrow 0$$

gives rise to an isomorphism  $\mathcal{L} \otimes \mathcal{M} \cong \bigwedge^2 \mathcal{V}$ . But  $\bigwedge^2 \mathcal{V}$  is the vector bundle underlying the variation of Hodge structure  $\bigwedge^2 \mathbb{V}$ , which is canonically isomorphic to the constant variation  $\mathbb{R}(-1)$  via the polarisation.  $\square$

Recall that the connection  $\nabla: \mathcal{V} \rightarrow \mathcal{V} \otimes \Omega_X^1$  is  $\mathbb{C}$ -linear and satisfies the Leibniz rule:

$$\nabla(fv) = f\nabla(v) + v \otimes df \quad (2)$$

for sections  $f \in \mathcal{O}_X(U)$ ,  $v \in \mathcal{V}(U)$ . We then define a Higgs field  $\vartheta: \mathcal{L} \rightarrow \mathcal{L}^{-1} \otimes_{\mathcal{O}_X} \Omega_X^1$  as the composition of inclusion, connection and projection:

$$\vartheta: \mathcal{L} \hookrightarrow \mathcal{V} \xrightarrow{\nabla} \mathcal{V} \otimes \Omega_X^1 \rightarrow \mathcal{M} \otimes \Omega_X^1 \cong \mathcal{L}^{-1} \otimes \Omega_X^1.$$

This is indeed an  $\mathcal{O}_X$ -linear homomorphism of sheaves: if  $v$  is a section of  $\mathcal{L}$ , then the second summand in (2) vanishes modulo  $\mathcal{L} \otimes \Omega_X^1$ .

Similarly, if  $\bar{X}$  is a smooth compact complex manifold,  $Y \subset \bar{X}$  is a normal crossings divisor,  $X = \bar{X} \setminus Y$  and  $\mathbb{V}$  is a real variation of elliptic curve type on  $X$  with unipotent monodromy around  $Y$ , then the Deligne extension of the corresponding vector bundles gives rise to a logarithmic Higgs bundle  $\vartheta: \mathcal{L} \rightarrow \mathcal{L}^{-1} \otimes \Omega_{\bar{X}}^1(\log Y)$ . If  $\bar{X}$  is a complex projective variety then  $\mathcal{L}$  and  $\vartheta$  are again algebraic by ‘GAGA’.

**2.3. Theta characteristics.** Theta characteristics are structures on algebraic curves or Riemann surfaces that can be expressed in many different ways. They also play a rôle in our theory, in the following two guises:

- (i) Let  $C$  be a smooth curve over a field  $k$ , let  $\bar{C}$  be its smooth projective completion and let  $D = \bar{C} \setminus C$  be the boundary. A (logarithmic) theta characteristic on  $C$  is an algebraic line bundle  $\mathcal{L}$  on  $\bar{C}$  such that  $\mathcal{L}^{\otimes 2} \simeq \omega_{\bar{C}}(\log D)$ . Let  $\Theta(C)$  be the set of isomorphism classes of logarithmic theta characteristics on  $C$ .

Assume from now on that  $k$  is algebraically closed of characteristic different from two. From the structure of the Picard group  $\text{Pic } \bar{C}$  one easily derives that  $\Theta(C)$  is nonempty if and only if the cardinality of  $D$  is even,

and in this case it has  $2^{2g}$  elements where  $g$  is the genus of  $\overline{C}$ . More precisely, consider the long exact cohomology sequence for the short exact sequence

$$1 \rightarrow \mu_2 \rightarrow \mathcal{O}_{\overline{C}}^\times \xrightarrow{(\cdot)^2} \mathcal{O}_{\overline{C}}^\times \rightarrow 1 \quad (3)$$

of étale sheaves on  $\overline{C}$ ; a segment of it can be rewritten as

$$\cdots k^\times \xrightarrow{(\cdot)^2} k^\times \xrightarrow{0} H_{\text{ét}}^1(\overline{C}, \mu_2) \rightarrow \text{Pic } \overline{C} \xrightarrow{2} \text{Pic } \overline{C} \rightarrow H_{\text{ét}}^2(\overline{C}, \mu_2) \rightarrow 0$$

(cf. the discussion in [10, III.3]). So,  $H_{\text{ét}}^1(\overline{C}, \mu_2)$ , which is a  $(\mathbb{Z}/2\mathbb{Z})$ -vector space of rank  $2g$ , is identified with the 2-torsion subgroup of  $\text{Pic } \overline{C}$ , and hence  $\Theta(C)$  is, when nonempty, a torsor under  $H_{\text{ét}}^1(\overline{C}, \mu_2)$ .

The same discussion can be repeated for Riemann surfaces of finite type, with holomorphic line bundles instead of algebraic line bundles and sheaves for the analytic topology instead of the étale topology.

- (ii) Let  $\Delta \subset \text{PGL}_2(\mathbb{R})^+ \cong \text{SL}_2(\mathbb{R})/\{\pm 1\}$  be a torsion-free lattice. Then a *theta characteristic for  $\Delta$*  is a lift of  $\Delta$  to  $\text{SL}_2(\mathbb{R})$ , i.e. a group homomorphism  $\Delta \rightarrow \text{SL}_2(\mathbb{R})$  such that the composition  $\Delta \rightarrow \text{SL}_2(\mathbb{R}) \rightarrow \text{PGL}_2(\mathbb{R})^+$  is the given embedding. A theta characteristic of  $\Delta$  is called *unipotent* if all parabolic elements of  $\Delta$  are lifted to unipotent elements of  $\text{SL}_2(\mathbb{R})$ . Denote the set of theta characteristics of  $\Delta$  by  $\Theta(\Delta)$ , and the subset of unipotent ones by  $\Theta^{\text{uni}}(\Delta) \subseteq \Theta(\Delta)$ .

Again it is a classical result, rediscovered several times, that there always exists a theta characteristic for  $\Delta$ : see [6, 17, 24, 43]. A unipotent theta characteristic exists if and only if the number of cusps for  $\Delta$  is even, cf. the discussion preceding [17, Theorem 4]. If  $\varrho: \Delta \rightarrow \text{SL}_2(\mathbb{R})$  is a theta characteristic, then every other theta characteristic of  $\Delta$  can be obtained as  $\varrho \cdot \varepsilon$ , where  $\varepsilon$  runs through all group homomorphisms  $\Delta \rightarrow \mu_2 = \{\pm 1\}$ . Hence  $\Theta(\Delta)$  is a torsor under the group  $\text{Hom}(\Delta, \mu_2) = H^1(\Delta, \mu_2)$ . Similarly  $\Theta^{\text{uni}}(\Delta)$ , if nonempty, is a torsor under the group  $H_{\text{cusp}}^1(\Delta, \mu_2)$  of all group homomorphisms  $\Delta \rightarrow \mu_2$  sending all parabolic elements to 1.

It is no coincidence that these two sets have such similar cohomological descriptions: if  $C$  is a smooth curve over  $\mathbb{C}$  and  $\Delta = \pi_1(C^{\text{an}})$  is viewed as a lattice in  $\text{PGL}_2(\mathbb{R})^+$  via the uniformisation theorem, then there is a natural bijection  $\Theta^{\text{uni}}(\Delta) \rightarrow \Theta(C)$ . It is equivariant for the isomorphism  $H_{\text{cusp}}^1(\Delta, \mu_2) \cong H_{\text{ét}}^1(\overline{C}, \mu_2)$  obtained from the identification of  $\pi_1(\overline{C}^{\text{an}})$  with the quotient of  $\Delta$  by the normal subgroup generated by its parabolic elements.

To construct it, let  $\varrho: \Delta \rightarrow \text{SL}_2(\mathbb{R})$  be an element of  $\Theta^{\text{uni}}(\Delta)$ , and let  $\Gamma = \varrho(\Delta)$ . Then  $C^{\text{an}} \cong \Gamma \backslash \mathbb{H}$ , and so we can view the Fuchsian variation  $\mathbb{W}_\Gamma$  as a variation on  $C^{\text{an}}$ . By construction it has unipotent monodromy around the cusps. Therefore we obtain a Higgs field  $\vartheta: \mathcal{L} \rightarrow \mathcal{L}^{-1} \otimes \Omega_{\overline{C}}^1(\log D)$ .

**Proposition 2.7.** *The Higgs field  $\vartheta: \mathcal{L} \rightarrow \mathcal{L}^{-1} \otimes \omega_{\overline{C}}(\log D)$  is an isomorphism, hence  $\mathcal{L}$  is a logarithmic theta characteristic on  $C$ .*

*Proof.* The Higgs field corresponding to the universal variation  $\mathbb{W}$  on  $\mathbb{H}$  is easily seen to be an isomorphism, hence so is the restriction of  $\vartheta$  to  $C$ . It remains to check isomorphy at the cusps. As a local model consider the quotient variation

$\mathbb{W}_{\Gamma_\infty}$  on  $\Gamma_\infty \backslash \mathbb{H} \cong \mathbb{D} \setminus \{0\}$  for  $\Gamma_\infty$  the cyclic group generated by  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ . Its underlying vector bundle is isomorphic to  $\mathcal{V} = \mathcal{O}^2$  on  $\mathbb{D} \setminus \{0\}$  with Hodge filtration  $F^1 \mathcal{V} = 0 \oplus \mathcal{O}$  and connection

$$\nabla = d - \begin{pmatrix} 0 & \frac{dt}{t} \\ 0 & 0 \end{pmatrix}.$$

From this description one directly reads off the Deligne extension of  $\mathbb{W}_{\Gamma_\infty}$  to  $\mathbb{D}$  and its corresponding Higgs field, which is given by

$$\mathcal{O}_{\mathbb{D}} \rightarrow \mathcal{O}_{\mathbb{D}} \otimes \omega_{\mathbb{D}}(0), \quad f(t) \mapsto \frac{f(t) dt}{t},$$

hence an isomorphism.  $\square$

The map  $\Gamma \mapsto [\mathcal{L}]$  is the desired bijection  $\Theta^{\text{uni}}(\Delta) \rightarrow \Theta(C)$ . Its equivariance for the natural isomorphism  $H_{\text{cusp}}^1(\Delta, \mu_2) \rightarrow H_{\text{et}}^1(\overline{C}, \mu_2)$  can be checked by an elementary calculation involving automorphy factors, similar to the proof of [17, Theorem 1].

**Proposition 2.8.** *Let  $\overline{C}$  be a smooth projective curve over  $\mathbb{C}$ , let  $D \subset \overline{C}(\mathbb{C})$  be a finite subset and let  $C = \overline{C} \setminus D$ . Assume that  $C$  is of hyperbolic type, i.e. the universal covering of  $C$  is  $\mathbb{H}$ . Let  $\mathbb{V}$  be a real variation of elliptic curve type on  $C^{\text{an}}$  which has unipotent monodromy around all cusps and let  $\vartheta: \mathcal{L} \rightarrow \mathcal{L}^{-1} \otimes \omega_{\overline{C}}(\log D)$  be the associated Higgs field on  $\overline{C}$ .*

*Then the following are equivalent:*

- (i) *The monodromy representation  $\pi_1(C^{\text{an}}) \rightarrow \text{SL}_2(\mathbb{R})$  of  $\mathbb{V}_{\mathbb{R}}$  is conjugate to a unipotent theta characteristic of  $\pi_1(C^{\text{an}})$ .*
- (ii) *The period map  $p: \widehat{C^{\text{an}}} \rightarrow \mathbb{H}$  is a biholomorphism.*
- (iii)  *$\vartheta$  is an isomorphism of line bundles ('the Higgs field is maximal').*
- (iv)  *$\mathcal{L}$  is a logarithmic theta characteristic on  $C$ .*
- (v) *The degree of the line bundle  $\mathcal{L}$  is  $-\frac{1}{2}\chi(C)$ , where  $\chi$  denotes the topological Euler characteristic.*

*If these conditions are satisfied, then the unipotent theta characteristic of  $\pi_1(C^{\text{an}})$  in (i) and the theta characteristic of  $C$  in (iv) correspond to each other under the bijection  $\Theta^{\text{uni}}(\pi_1(C^{\text{an}})) \leftrightarrow \Theta(C)$ .*

*Proof.* The equivalence (i)  $\Leftrightarrow$  (ii) is an elementary consequence of the fact that the period map is equivariant for the monodromy representation, and (ii)  $\Leftrightarrow$  (iii) is [53, Lemma 2.1]. A more abstract version of this argument can be found in the discussion preceding [48, Prop. 9.1]).

The implications (iii)  $\Rightarrow$  (iv)  $\Rightarrow$  (v) are obvious. For the remaining part (v)  $\Rightarrow$  (iii) note that if (v) is true then  $\vartheta$  cannot be identically zero: assume it is. Then by construction the Hodge decomposition is flat, i.e.  $\mathcal{L} = F^1 \mathcal{V}$  is the line bundle underlying a local subsystem of rank one in  $\mathbb{V}_{\mathbb{C}}$ . But such a line bundle has degree zero, contradiction. Hence the Higgs field is a nontrivial holomorphic section of the degree zero line bundle  $\underline{\text{Hom}}(\mathcal{L}, \mathcal{L}^{-1} \otimes \omega_{\overline{C}}(\log D))$ ; but as the degree is the number of zeros minus the number of poles and  $\vartheta$  has no poles, it cannot have any zeros either, hence it is invertible.  $\square$

### 3. QUATERNIONIC SHIMURA VARIETIES

We now review the construction and some important properties of the Shimura varieties that occur as targets for modular embeddings. They have been studied extensively in [47] in rather classical language; however, for our approach the adelic formalism developed by Deligne [9, 11] is more suitable. We state all definitions necessary for our setup, but for their incorporation into a general theory the reader is advised to consult [9, 11, 34, 35].

**3.1. The Shimura varieties themselves.** Let  $F$  be a totally real number field and let  $\mathcal{S} = \mathcal{S}_f \cup \mathcal{S}_\infty$  be its set of places. Let  $B$  be a quaternion algebra over  $F$  and let  $\mathcal{R}(B) = \mathcal{R}_f(B) \cup \mathcal{R}_\infty(B)$  be the set of places where  $B$  is ramified. Throughout we assume Eichler's condition: the set  $\mathcal{P} = \mathcal{S}_\infty \setminus \mathcal{R}_\infty(B)$  is nonempty. For every  $\varrho \in \mathcal{P}$  we fix an isomorphism

$$B \otimes_{F, \varrho} \mathbb{R} \cong M_2(\mathbb{R}). \quad (4)$$

Let  $H$  be the linear algebraic group over  $F$  with  $H(F) = B^1$  (the group of invertible elements in  $B$  with reduced norm one); let  $G = \text{Res}_{F/\mathbb{Q}} H$ . Then (4) induces, for every  $\varrho \in \mathcal{P}$ , an isomorphism  $H_F \otimes_{F, \varrho} \mathbb{R} \cong \text{SL}_{2, \mathbb{R}}$ , and for every  $\varrho \in \mathcal{R}_\infty(B)$  we fix an isomorphism  $H_F \otimes_{F, \varrho} \mathbb{R} \cong \text{SU}_{2, \mathbb{R}}$ . These glue to an isomorphism

$$G_{\mathbb{R}} \cong \prod_{\varrho \in \mathcal{S}_\infty} H_F \otimes_{F, \varrho} \mathbb{R} \rightarrow \prod_{\varrho \in \mathcal{P}} \text{SL}_{2, \mathbb{R}} \times \prod_{\varrho \in \mathcal{R}_\infty(B)} \text{SU}_{2, \mathbb{R}}. \quad (5)$$

Let  $G^{\text{ad}}$  be the adjoint group of  $G$ , and similarly for  $H$ , so that  $G^{\text{ad}}(\mathbb{Q}) = H^{\text{ad}}(F) = B^\times / F^\times$ . Then the isomorphism (5) induces an isomorphism

$$G_{\mathbb{R}}^{\text{ad}} \cong \prod_{\varrho \in \mathcal{S}_\infty} H_F^{\text{ad}} \otimes_{F, \varrho} \mathbb{R} \rightarrow \prod_{\varrho \in \mathcal{P}} \text{PGL}_{2, \mathbb{R}} \times \prod_{\varrho \in \mathcal{R}_\infty(B)} \text{PU}_{2, \mathbb{R}}. \quad (6)$$

Let  $\mathbb{S} = \text{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_m$  be the Deligne torus (cf. Section 2.1), and let  $h_1: \mathbb{S} \rightarrow \text{PGL}_{2, \mathbb{R}}$  be the standard homomorphism, given on  $\mathbb{R}$ -points by

$$\mathbb{C}^\times \rightarrow \text{PGL}_2(\mathbb{R}), \quad a + bi \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \bmod \mathbb{R}^\times.$$

Let  $h: \mathbb{S} \rightarrow G_{\mathbb{R}}^{\text{ad}}$  be the homomorphism which in the product decomposition (6) is given by  $h_1$  on each  $\text{PGL}_{2, \mathbb{R}}$ -factor and by the trivial homomorphism on each  $\text{PU}_{2, \mathbb{R}}$ -factor. Let  $X \subset \text{Hom}(\mathbb{S}, G_{\mathbb{R}}^{\text{ad}})$  be the  $G^{\text{ad}}(\mathbb{R})^+$ -conjugacy class containing  $h$ .

The pair  $(G, X)$  is a connected Shimura datum in the following sense:

**Definition 3.1.** *A connected Shimura datum is a pair  $(G, X)$  consisting of a semisimple algebraic group  $G$  over  $\mathbb{Q}$  and a  $G^{\text{ad}}(\mathbb{R})^+$ -conjugacy class  $X$  of homomorphisms  $h: \mathbb{S} \rightarrow G_{\mathbb{R}}^{\text{ad}}$  such that the following conditions are satisfied for one (and hence all)  $h \in X$ :*

- (i) *No characters other than 1,  $z/\bar{z}$  and  $\bar{z}/z$  occur in the representation of  $\mathbb{S}$  on  $\text{Lie } G_{\mathbb{C}}^{\text{ad}}$ .*
- (ii)  *$\text{ad } h(i)$  is a Cartan involution on  $G^{\text{ad}}$ .*
- (iii)  *$h$  projects nontrivially to every  $\mathbb{Q}$ -factor of  $G^{\text{ad}}$ .*

The definition given here is the ‘alternative definition’ [35, Definition 4.22]. It is also used in [33, p. 93] on which we heavily rely.

*Congruence subgroups.* Recall that an automorphism of a finite-dimensional  $\mathbb{C}$ -vector space is *neat* if its eigenvalues generate a torsion-free subgroup of  $\mathbb{C}^\times$ , and an arithmetic subgroup  $\Gamma \subset R(\mathbb{Q})$  of a linear algebraic group  $R$  over  $\mathbb{Q}$  is neat if the image of  $\Gamma$  under one (equivalently, any) faithful rational representation of  $R$  has only neat elements. Every sufficiently small congruence subgroup of  $R(\mathbb{Q})$  is neat by [1, 17.4]. Note that neat subgroups are torsion-free.

Let  $G$  be a semisimple linear algebraic group over  $\mathbb{Q}$ . The set of neat congruence subgroups of  $G(\mathbb{Q})$  is denoted by  $\tilde{\Sigma}(G)$ . The completion of  $G(\mathbb{Q})$  with respect to the topology induced by the subgroups in  $\tilde{\Sigma}(G)$  is canonically identified with  $G(\mathbb{A}^f)$ .

The set of neat arithmetic subgroups of  $G^{\text{ad}}(\mathbb{Q})^+$  that contain the image of a congruence subgroup of  $G(\mathbb{Q})$  is denoted by  $\Sigma(G)$ . Note that these are not necessarily congruence subgroups of  $G^{\text{ad}}(\mathbb{Q})^+$ . If  $\tilde{\Gamma} \in \tilde{\Sigma}(G)$  then its image in  $G^{\text{ad}}(\mathbb{Q})$  is an element of  $\Sigma(G)$  denoted by  $\text{ad } \tilde{\Gamma}$ . The completion of  $G^{\text{ad}}(\mathbb{Q})^+$  with respect to the topology induced by these subgroups is denoted by  $\mathcal{A}(G)$ . Note that this is the group denoted by  $G^{\text{ad}}(\mathbb{Q})^{+\wedge}$  in [11] and [33]. If  $G$  is simply connected (as in all cases we consider) the canonical maps  $\text{ad}: G(\mathbb{A}^f) \rightarrow \mathcal{A}(G)$  and  $G^{\text{ad}}(\mathbb{Q})^+ \rightarrow \mathcal{A}(G)$  define an isomorphism of topological groups

$$G(\mathbb{A}^f) *_{G(\mathbb{Q})} G^{\text{ad}}(\mathbb{Q})^+ \cong \mathcal{A}(G), \quad (7)$$

see [11, 2.1.6.2].

*Quotients.* The space  $X$  admits a natural structure as a complex manifold, in fact as an Hermitian symmetric domain. The operation of  $G^{\text{ad}}(\mathbb{R})^+$  on  $X$  by conjugation identifies  $G^{\text{ad}}(\mathbb{R})^+$  with the group of biholomorphisms  $X \rightarrow X$ . Each arithmetic subgroup  $\Gamma \subset G^{\text{ad}}(\mathbb{Q})^+$  containing a group in  $\Sigma(G)$  operates freely and properly discontinuously on  $X$ ; the quotient  $\Gamma \backslash X$  is a normal complex space, and by [2, Theorem 3.10] it admits a unique structure as a normal quasiprojective complex variety. We denote this variety by  $\text{Sh}_\Gamma^0(G, X)$ . If  $\Gamma \in \Sigma(G)$  this variety is smooth, and  $X$  is the universal cover of  $\Gamma \backslash X = \text{Sh}_\Gamma^0(G, X)(\mathbb{C})$ . The system consisting of the  $\text{Sh}_\Gamma^0(G, X)$  for varying  $\Gamma \in \Sigma(G)$  together with the obvious ‘forgetful’ maps  $\text{Sh}_{\Gamma'}^0(G, X) \rightarrow \text{Sh}_\Gamma^0(G, X)$  for  $\Gamma' \subseteq \Gamma$ , as well as its limit

$$\text{Sh}^0(G, X) = \varprojlim_{\Gamma \in \Sigma(G)} \text{Sh}_\Gamma^0(G, X) = \varprojlim_{\tilde{\Gamma} \in \tilde{\Sigma}(G)} \text{Sh}_{\text{ad } \tilde{\Gamma}}^0(G, X)$$

(limits in the category of  $\mathbb{C}$ -schemes), is called the *connected Shimura variety* associated with  $(G, X)$ . There is a natural left action of  $G^{\text{ad}}(\mathbb{Q})^+$  on  $\text{Sh}^0(G, X)$ , given on finite levels by

$$g: \Gamma \backslash X \rightarrow g\Gamma g^{-1} \backslash X, \quad [x] \mapsto [gx].$$

By continuity it extends to an action of  $\mathcal{A}(G)$  on  $\text{Sh}^0(G, X)$ . If  $\Gamma \in \Sigma(G)$  then the finite type variety  $\text{Sh}_\Gamma^0(G, X)$  can be retrieved as the quotient  $\hat{\Gamma} \backslash \text{Sh}^0(G, X)$  where  $\hat{\Gamma}$  is the topological closure of  $\Gamma$  in  $\mathcal{A}(G)$ .

*Adelic description.* Let  $\tilde{\Gamma} \in \tilde{\Sigma}(G)$  and let  $K$  be its closure in  $G(\mathbb{A}^f)$ . Letting  $G(\mathbb{Q})$  operate from the left and  $K$  from the right on  $X \times G(\mathbb{A}^f)$  by

$$q(x, g)k = (\text{ad}(q)x, qgk)$$

we can form the double quotient  $G(\mathbb{Q}) \backslash X \times G(\mathbb{A}^f)/K$ . From the general formalism of Shimura varieties it follows that the map

$$\mathrm{ad} \tilde{F} \backslash X \rightarrow G(\mathbb{Q}) \backslash X \times G(\mathbb{A}^f)/K, \quad [x] \mapsto [x, 1]$$

is biholomorphic. Passing to the limit over smaller and smaller  $\tilde{F}$  we obtain a bijection

$$\mathrm{Sh}^0(G, X)(\mathbb{C}) \cong G(\mathbb{Q}) \backslash X \times G(\mathbb{A}^f) \quad (8)$$

that can even be considered as a biholomorphism for suitable topologies and complex structures on both sides, see [11, Section 2.7]. The action of  $\mathcal{A}(G)$  on  $\mathrm{Sh}^0(G, X)$  translates via (7) and (8) to the following two left actions on  $G(\mathbb{Q}) \backslash X \times G(\mathbb{A}^f)$ :

- (i)  $G(\mathbb{A}^f)$  operates by  $a: [x, g] \mapsto [x, ga^{-1}]$ ;
- (ii)  $G^{\mathrm{ad}}(\mathbb{Q})^+$  operates by  $q: [x, g] \mapsto [qx, \mathrm{ad}(q)(g)]$ .

Since these two actions agree on  $G(\mathbb{Q})$ , they indeed glue to an action by  $\mathcal{A}(G)$ .

*Dichotomy: projective or modular.* Let  $\Gamma \in \Sigma(G)$  and let  $S = \mathrm{Sh}_\Gamma^0(G, X)$ . Then  $S$  is projective if and only if  $B$  is a division algebra; this is easily seen using the compactness criterion [3, Theorem 11.6] of Borel and Harish-Chandra. Otherwise  $B$  has to be isomorphic to  $M_2(F)$ , hence  $G$  is isomorphic to  $\mathrm{Res}_{F/\mathbb{Q}} \mathrm{SL}_2$ . The associated Shimura varieties are then called *Hilbert–Blumenthal* varieties and can be interpreted as moduli spaces for abelian varieties  $A$  of dimension  $[F : \mathbb{Q}]$  with a ring homomorphism  $F \rightarrow \mathrm{End}(A) \otimes \mathbb{Q}$ , together with certain level structures whose exact form depends on  $\Gamma$ . For more about these varieties see [52].

*The compact dual Hermitian symmetric space.* For each  $h: \mathbb{S} \rightarrow G_{\mathbb{R}}^{\mathrm{ad}}$  in  $X$  we can consider the associated Hodge filtration homomorphism  $\mu_h: \mathbb{G}_{m, \mathbb{C}} \rightarrow G_{\mathbb{C}}^{\mathrm{ad}}$ . The  $G^{\mathrm{ad}}(\mathbb{C})$ -conjugacy class  $\check{X} \subset \mathrm{Hom}(\mathbb{G}_{m, \mathbb{C}}, G_{\mathbb{C}}^{\mathrm{ad}})$  containing  $\mu_h$  depends only on  $X$  and not on the chosen representative  $h \in X$ . The space  $\check{X}$  has a natural structure as a compact complex manifold, and in fact as a complex algebraic variety. The

*Borel embedding*

$$\beta: X \hookrightarrow \check{X}, \quad h \mapsto \mu_h,$$

realises  $X$  as an open complex submanifold of  $\check{X}$ .

*Classical description.* We let  $G^{\mathrm{ad}}(\mathbb{R})^+$  act on  $\mathbb{H}^{\mathcal{P}}$  by coordinate-wise Möbius transformations via the projection

$$G^{\mathrm{ad}}(\mathbb{R})^+ \stackrel{(6)}{\cong} \prod_{\varrho \in \mathcal{P}} \mathrm{PGL}_2(\mathbb{R})^+ \times \prod_{\varrho \in \mathcal{R}_\infty(B)} \mathrm{PU}_2(\mathbb{R}) \twoheadrightarrow \prod_{\varrho \in \mathcal{P}} \mathrm{PGL}_2(\mathbb{R})^+.$$

Then for each  $h \in X$  the image  $h(\mathbb{C}^\times) \subset G^{\mathrm{ad}}(\mathbb{R})^+$  has a unique common fixed point in  $\mathbb{H}^{\mathcal{P}}$ , and sending  $h$  to that fixed point defines a biholomorphism  $X \rightarrow \mathbb{H}^{\mathcal{P}}$ .

Similarly, the identification (6) defines a projection

$$G^{\mathrm{ad}}(\mathbb{C}) \stackrel{(6)}{\cong} \prod_{\varrho \in \mathcal{P}} \mathrm{PGL}_2(\mathbb{C}) \times \prod_{\varrho \in \mathcal{R}_\infty(B)} \mathrm{PU}_2(\mathbb{C}) \twoheadrightarrow \prod_{\varrho \in \mathcal{P}} \mathrm{PGL}_2(\mathbb{C}).$$

We let  $G^{\mathrm{ad}}(\mathbb{C})$  operate via this projection and by coordinate-wise Möbius transformations on  $\mathbb{P}^1(\mathbb{C})^{\mathcal{P}}$ . For each  $\mu: \mathbb{G}_m \rightarrow G_{\mathbb{C}}^{\mathrm{ad}}$  in  $\check{X}$  the image  $\mu(\mathbb{C}) \subset G^{\mathrm{ad}}(\mathbb{C})$

has a unique fixed point in  $\mathbb{P}^1(\mathbb{C})^{\mathcal{P}}$ , and sending  $\mu$  to that fixed point defines a biholomorphism  $\check{X} \rightarrow \mathbb{P}^1(\mathbb{C})^{\mathcal{P}}$ .

The maps constructed in this paragraph fit in a commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{\beta} & \check{X} \\ \downarrow \simeq & & \downarrow \simeq \\ \mathbb{H}^{\mathcal{P}} & \hookrightarrow & \mathbb{P}^1(\mathbb{C})^{\mathcal{P}} \end{array}$$

with the natural inclusion on the lower horizontal line; the morphisms in this diagram are equivariant for the corresponding morphisms in the commutative diagram

$$\begin{array}{ccc} G^{\text{ad}}(\mathbb{R})^+ & \hookrightarrow & G^{\text{ad}}(\mathbb{C}) \\ \downarrow \simeq & & \downarrow \simeq \\ (\text{PGL}_2(\mathbb{R})^+)^{\mathcal{P}} & \hookrightarrow & \text{PGL}_2(\mathbb{C})^{\mathcal{P}}. \end{array}$$

**3.2. Automorphic line bundles.** Let  $\mathfrak{L} \rightarrow \mathbb{P}^1$  be the total space of the line bundle  $\mathcal{O}(1)$  on  $\mathbb{P}^1$ . Recall that the fibre  $\mathfrak{L}_x$  over a point  $x \in \mathbb{P}^1(\mathbb{C})$  represented by a one-dimensional subspace  $W \subset \mathbb{C}^2$  is the dual space  $W^* = \text{Hom}(W, \mathbb{C})$ . Therefore the points of  $\mathfrak{L}(\mathbb{C})$  are pairs  $(W, \lambda)$  where  $W \subset \mathbb{C}^2$  is a line and  $\lambda: W \rightarrow \mathbb{C}$  is a linear form. Then

$$\gamma \cdot (W, \lambda) \stackrel{\text{def}}{=} (\gamma(W), \lambda \circ \gamma^{-1})$$

defines an action of  $\text{SL}_2(\mathbb{C})$  on  $\mathfrak{L}(\mathbb{C})$  which is fact an algebraic action of  $\text{SL}_2$  on  $\mathfrak{L}$  defined over  $\mathbb{Q}$  and compatible with the vector bundle structure on  $\mathfrak{L}$ .

For each embedding  $\varrho \in \mathcal{P}$  let  $\mathfrak{L}_{\varrho}$  be the line bundle on  $\check{X}$  which is the pullback of  $\mathfrak{L}$  along the composition

$$\check{X} \cong \mathbb{P}^1(\mathbb{C})^{\mathcal{P}} \xrightarrow{\text{pr}_{\varrho}} \mathbb{P}^1(\mathbb{C}).$$

Since this holomorphic map is equivariant for the corresponding projection on groups,

$$G(\mathbb{C}) \twoheadrightarrow \text{SL}_2(\mathbb{C})^{\mathcal{P}} \xrightarrow{\text{pr}_{\varrho}} \text{SL}_2(\mathbb{C}),$$

the line bundle  $\mathfrak{L}_{\varrho} \rightarrow \check{X}$  becomes a  $G_{\mathbb{C}}$ -vector bundle in a natural way. That is,  $G_{\mathbb{C}}$  operates on  $\mathfrak{L}_{\varrho}$  as an algebraic variety in a way which is compatible with the structure map  $\mathfrak{L}_{\varrho} \rightarrow \check{X}$  and which respects the vector space structures on the fibres.

This action restricts to a holomorphic action of  $G(\mathbb{R})$  on  $\mathfrak{L}_{\varrho}|_X \rightarrow X$ . For every subgroup  $\tilde{\Gamma} \in \tilde{\Sigma}(G)$  this vector bundle descends to a holomorphic vector bundle

$$L_{\varrho, \tilde{\Gamma}} \stackrel{\text{def}}{=} \tilde{\Gamma} \backslash (\mathfrak{L}_{\varrho}|_X) \rightarrow \text{ad } \tilde{\Gamma} \backslash X = \text{Sh}_{\text{ad } \tilde{\Gamma}}^0(G, X)(\mathbb{C}).$$

**Lemma 3.2.** *Assume that  $|\mathcal{P}| = \dim S > 1$ . Then there exists a unique algebraic structure  $\mathcal{L}_{\varrho, \tilde{\Gamma}}$  on the bundle  $L_{\varrho, \tilde{\Gamma}}$  compatible with the given holomorphic structure. Furthermore, any holomorphic global section of  $(\mathcal{L}_{\varrho, \tilde{\Gamma}})^{\otimes n} \otimes \Omega_S^p$  is algebraic.*

*Proof.* This is contained in [34, III. Lemma 2.2]. □

Also in the one-dimensional case there exists a canonical (but no more unique) algebraic line bundle underlying  $L_{\varrho, \tilde{\Gamma}}$  (see [33, Section 4] for details); we denote its sheaf of sections by  $\mathcal{L}_{\varrho, \tilde{\Gamma}}$  or, if  $\tilde{\Gamma}$  is clear from the context, simply by  $\mathcal{L}_{\varrho}$ . Furthermore, the line bundles  $\mathcal{L}_{\varrho, \tilde{\Gamma}}$  for varying  $\tilde{\Gamma}$  are compatible in the obvious way, hence we obtain a line bundle  $\mathcal{L}_{\varrho}$  on  $\mathrm{Sh}^0(G, X)$  together with a continuous action of  $G(\mathbb{A}^f)$  lifting that on  $\mathrm{Sh}^0(G, X)$ . Note this does not extend to an action of  $\mathcal{A}(G)$  on  $\mathcal{L}_{\varrho}$ .

*Remark 3.3.* The bundle  $\mathcal{L}_{\varrho, \tilde{\Gamma}}$  is the holomorphic line bundle over  $\mathrm{ad} \tilde{\Gamma} \backslash X$  with automorphy factors

$$j_{\varrho}(\tilde{\gamma}, z) = c_{\varrho} z_{\varrho} + d_{\varrho}$$

where  $z = (z_{\varrho})_{\varrho} \in \mathbb{H}^p$  and

$$\tilde{\gamma} \mapsto \begin{pmatrix} a_{\varrho} & b_{\varrho} \\ c_{\varrho} & d_{\varrho} \end{pmatrix} \quad (9)$$

under the  $\varrho$ -th factor projection in (5). That is, holomorphic sections of  $\mathcal{L}_{\varrho}$  over an open subset  $U \subseteq S(\mathbb{C})$  are the same as holomorphic functions  $f: V \rightarrow \mathbb{C}$ , where  $V$  is the preimage of  $U$  in  $X$ , such that

$$f(\tilde{\gamma}z) = j_{\varrho}(\tilde{\gamma}, z)f(z)$$

for all  $\tilde{\gamma} \in \tilde{\Gamma}$  and  $z \in V$ . Hence global sections of  $\mathcal{L}_{\varrho, \tilde{\Gamma}}$  can be identified with holomorphic modular forms for  $\tilde{\Gamma}$  of weight one in direction  $\varrho$  and weight zero in all other directions.

*Remark 3.4.* The reason why we work with connected Shimura varieties instead of the more customary nonconnected Shimura varieties defined by the groups  $G_0$  with  $G_0(\mathbb{Q}) = B^{\times}$  is that the theory of automorphic vector bundles is much simpler in the connected case, with a necessity to use stacks in the nonconnected case, cf. [34, Section III.8]. In our situation, any congruence subgroup of  $G_0(\mathbb{Q}) = B^{\times}$  contains a finite index subgroup of  $\mathfrak{o}_K^{\times}$  and hence elements that operate trivially on  $X$  but not on  $\mathcal{L}_{\varrho}$ .

Now fix some subgroup  $\tilde{\Gamma} \in \tilde{\Sigma}(G)$  with image  $\Gamma = \mathrm{ad} \tilde{\Gamma} \in \Sigma(G)$  and set  $S = \mathrm{Sh}_{\Gamma}^0(G, X)$ . The line bundles  $\mathcal{L}_{\varrho} = \mathcal{L}_{\varrho, \tilde{\Gamma}}$  on  $S$  are related to the cotangent bundle  $\Omega_S^1$ . Recall that the cotangent bundle of  $\mathbb{P}^1$  is *canonically* isomorphic to  $\mathcal{L}^{\otimes 2}$ . Hence the cotangent bundle of  $X \simeq \mathbb{P}^1(\mathbb{C})^p$  is canonically isomorphic to

$$\bigoplus_{\varrho \in \mathcal{P}} \mathcal{L}_{\varrho}^{\otimes 2}$$

as a  $G_{\mathbb{C}}$ -vector bundle. This in turn implies that there is a canonical isomorphism of algebraic vector bundles on  $S$ :

$$\alpha: \bigoplus_{\varrho \in \mathcal{P}} \mathcal{L}_{\varrho}^{\otimes 2} \cong \Omega_S^1.$$

From this we can easily construct a Higgs field  $\vartheta_{\varrho}: \mathcal{L}_{\varrho} \rightarrow \mathcal{L}_{\varrho}^{-1} \otimes \Omega_S^1$  as the composition

$$\vartheta_{\varrho}: \mathcal{L}_{\varrho} \cong \mathcal{L}_{\varrho}^{-1} \otimes \mathcal{L}_{\varrho}^{\otimes 2} \hookrightarrow \mathcal{L}_{\varrho}^{-1} \otimes \bigoplus_{\sigma \in \mathcal{P}} \mathcal{L}_{\sigma}^{\otimes 2} \xrightarrow{\mathrm{id} \otimes \alpha} \mathcal{L}_{\varrho}^{-1} \otimes \Omega_S^1.$$

These Higgs fields come from variations of Hodge structure:

For every  $\varrho \in \mathcal{P}$  let  $\mathbb{W}_\varrho$  be the pullback of the universal variation  $\mathbb{W}$  on  $\mathbb{H}$  along the projection  $\text{pr}_\varrho: \mathbb{H}^\mathcal{P} \rightarrow \mathbb{H}$ . I.e., it is the polarised variation of real Hodge structure on  $\mathbb{H}^\mathcal{P}$  whose underlying local system is the trivial local system  $W = \mathbb{R}^2$  with the symplectic form  $\psi$  as in (1), and whose Hodge decomposition is of the form

$$(\mathbb{W}_\varrho)_{z, \mathbb{C}} = W_z^{1,0} \oplus W_z^{0,1}, \quad W_z^{1,0} = \mathbb{C} \cdot \begin{pmatrix} z_\varrho \\ 1 \end{pmatrix}$$

(here  $z = (\dots, z_\varrho, \dots) \in \mathbb{H}^\mathcal{P}$ ).

In complete analogy to the one-dimensional case, the operation of  $\text{SL}_2(\mathbb{R})^\mathcal{P}$  by coordinate-wise Möbius transformations extends to an action on  $\mathbb{W}_\varrho$ . For this reason,  $\mathbb{W}_\varrho$  can be interpreted as a real variation of elliptic curve type on  $X$  with an action of  $G(\mathbb{R})$ . It therefore descends to a variation  $\mathbb{W}_{\varrho, \tilde{\Gamma}}$  on  $S$  whose monodromy representation is given by (9).

**Lemma 3.5.** *Suppress all subscripts  $\tilde{\Gamma}$  for better readability. There is a canonical isomorphism of holomorphic line bundles  $F^1 \mathcal{W}_\varrho \rightarrow \mathcal{L}_\varrho^{\text{an}}$  such that the diagram*

$$\begin{array}{ccc} F^1 \mathcal{W}_\varrho & \xrightarrow{\eta} & (F^1 \mathcal{W}_\varrho)^{-1} \otimes \Omega_S^1 \\ \downarrow \simeq & & \downarrow \simeq \\ \mathcal{L}_\varrho^{\text{an}} & \xrightarrow{j_\varrho^{\text{an}}} & \mathcal{L}_\varrho^{-1, \text{an}} \otimes \Omega_S^1 \end{array}$$

*commutes up to a sign. Here  $\eta$  is the Higgs field associated with the variation of Hodge structure  $\mathbb{W}_\varrho$ .*

*Proof.* These isomorphisms can be realised as follows: a lift of  $\mathcal{L}_\varrho^{\text{an}} \rightarrow F^1 \mathcal{W}_\varrho$  to  $X$  sends an automorphic function  $f$  as in Remark 3.3 to the section  $\begin{pmatrix} z_\varrho f(z) \\ f(z) \end{pmatrix}$  of  $F^1 \mathcal{W}_\varrho \subset \mathcal{W}_\varrho$ , and a lift of  $\mathcal{W}_\varrho / F^1 \mathcal{W}_\varrho \rightarrow \mathcal{L}_\varrho^{-1, \text{an}}$  to  $X$  sends the equivalence class of  $\begin{pmatrix} f(z) \\ g(z) \end{pmatrix}$  to the function  $z_\varrho g(z) - f(z)$  which is automorphic for the system of automorphy factors  $j_\varrho(\tilde{\gamma}, z)^{-1}$ .  $\square$

*Hilbert–Blumenthal case.* These objects have a more geometric description in the case where  $B = \text{M}_2(F)$ . Recall that then for any  $\Gamma = \text{ad} \tilde{\Gamma}$  with  $\tilde{\Gamma} \in \tilde{\Sigma}(G)$  the variety  $S = \text{Sh}_\Gamma^0(G, X)$  has an interpretation as a fine moduli space for abelian varieties with certain PEL structures, in particular there is a universal family  $p: A \rightarrow S$  of polarised abelian varieties of rank  $[F : \mathbb{Q}]$  together with a ring homomorphism  $F \hookrightarrow \text{End}(A/S) \otimes \mathbb{Q}$ . The polarised variation  $R^1 p_* \mathbb{R}$  decomposes as a direct sum

$$R^1 p_* \mathbb{R} = \bigoplus_{\varrho: F \rightarrow \mathbb{R}} (R^1 p_* \mathbb{R})_\varrho$$

where  $a \in F$  acts on the summand with subscript  $\varrho$  as multiplication by  $\varrho(a)$ . A direct calculation shows that

$$R^1 p_* \mathbb{R} \cong \mathbb{W}_\varrho. \quad (10)$$

The holomorphic vector bundle underlying  $R^1 p_* \mathbb{C}$  is canonically isomorphic to the relative de Rham cohomology

$$\mathcal{H}_{\text{dR}}^1(A/S) = R^1 p_*(\Omega_{A/S}^\bullet)$$

with the Hodge filtration given by the short exact sequence

$$0 \rightarrow p_*\Omega_{A/S}^1 \rightarrow \mathcal{H}_{\text{dR}}^1(A/S) \rightarrow R^1p_*\mathcal{O}_A \rightarrow 0 \quad (11)$$

and the Gauss–Manin connection  $\nabla^{\text{GM}}: \mathcal{H}_{\text{dR}}^1(A/S) \rightarrow \mathcal{H}_{\text{dR}}^1(A/S) \otimes \Omega_{S/\mathbb{C}}^1$ . All these structures respect the  $F$ -action on cohomology and hence decompose into eigenspaces.

**Proposition 3.6.** *Let  $S$  be a Hilbert–Blumenthal variety with universal family  $p: A \rightarrow S$ . Let  $\mathcal{H}_\varrho$  be the subbundle of  $\mathcal{H}_{\text{dR}}^1(A/S)$  where  $F$  operates through  $\varrho$ , similarly for  $(\Omega_{A/S}^1)_\varrho$  and  $(R^1p_*\mathcal{O}_A)_\varrho$ . Then (11) restricts to a short exact sequence*

$$0 \rightarrow p_*(\Omega_{A/S}^1)_\varrho \rightarrow \mathcal{H}_\varrho \rightarrow (R^1p_*\mathcal{O}_A)_\varrho \rightarrow 0$$

and the Gauss–Manin connection restricts to a connection

$$\nabla_\varrho^{\text{GM}}: \mathcal{H}_\varrho \rightarrow \mathcal{H}_\varrho \otimes \Omega_S^1.$$

Furthermore there exist isomorphisms of algebraic line bundles

$$\mathcal{L}_\varrho \cong p_*(\Omega_{A/S}^1)_\varrho \quad \text{and} \quad \mathcal{L}_\varrho^{-1} \cong (R^1p_*\mathcal{O}_A)_\varrho$$

sending the Higgs field  $\vartheta_\varrho: \mathcal{L}_\varrho \rightarrow \mathcal{L}_\varrho^{-1} \otimes \Omega_S^1$  to the negative of the Higgs field associated with the connection  $\nabla_\varrho^{\text{GM}}$ .

*Proof.* If  $\dim S > 1$  (equivalently, if  $F \neq \mathbb{Q}$ ) the existence of these isomorphisms only needs to be checked in the analytic category by Lemma 3.2; this is accomplished by combining Lemma 3.5 with (10). The remaining case  $F = \mathbb{Q}$  is easy.  $\square$

#### 4. GEOMETRY OF MODULAR EMBEDDINGS

In this section we study regular maps from algebraic curves to quaternionic Shimura varieties as in the previous section.

**4.1. Modular embeddings.** Let  $(G, X)$  be a quaternionic connected Shimura datum as before, associated with a quaternion algebra  $B$  over a totally real number field  $F$ . We shall consider triples  $(\Delta, \varphi, \tilde{f})$  of the following form:

- (i)  $\Delta$  is a lattice in  $\text{PGL}_2(\mathbb{R})^+$ ;
- (ii)  $\varphi$  is a group homomorphism  $\varphi: \Delta \rightarrow G^{\text{ad}}(\mathbb{Q})^+$  whose image is contained in an arithmetic subgroup of  $G^{\text{ad}}(\mathbb{Q})^+$ ;
- (iii)  $\tilde{f}$  is a holomorphic map  $\tilde{f}: \mathbb{H} \rightarrow X$  which is equivariant for  $\varphi$ , i.e. which satisfies

$$\tilde{f}(\delta z) = \varphi(\delta)\tilde{f}(z) \quad \text{for all } \delta \in \Delta, z \in \mathbb{H}.$$

Using the classical description of  $(G, X)$  we can interpret  $\tilde{f}$  as a map  $\mathbb{H} \rightarrow \mathbb{H}^{\mathcal{P}}$ ; we denote the coordinate function  $\mathbb{H} \rightarrow \mathbb{H}$  for  $\varrho \in \mathcal{P}$  by  $\tilde{f}_\varrho$ . Similarly we can view  $\varphi$  as a homomorphism  $\Delta \rightarrow G^{\text{ad}}(\mathbb{R})^+$ , and via (6) we obtain for every  $\varrho \in \mathcal{P}$  a homomorphism  $\varphi_\varrho: \Delta \rightarrow \text{PGL}_2(\mathbb{R})^+$ .

**Proposition and Definition 4.1.** *Let  $(G, X)$  be a quaternionic Shimura datum associated with  $B/F$ , let  $(\Delta, \varphi, \tilde{f})$  as above, and let  $\varrho \in \mathcal{P}$ . The following are equivalent:*

- (i)  $\varphi_\varrho: \Delta \rightarrow \text{PGL}_2(\mathbb{R})^+$  is conjugate by an element of  $\text{PGL}_2(\mathbb{R})^+$  to the identity embedding  $\Delta \hookrightarrow \text{PGL}_2(\mathbb{R})^+$ .

(ii)  $\tilde{f}_\varrho: \mathbb{H} \rightarrow \mathbb{H}$  is a biholomorphism.

If this is the case we call  $(\varphi, \tilde{f})$  a modular embedding for  $\Delta$  with respect to  $\varrho$ .

*Proof.* The equivalence of the two conditions is a simple consequence of the Schwarz–Pick Lemma.  $\square$

If  $(\varphi, \tilde{f})$  is a modular embedding for  $\Delta$  we can adjust the isomorphism (6), whose precise choice is not essential for any of our constructions, in such a way that  $\varphi_\varrho: \Delta \rightarrow \mathrm{PGL}_2(\mathbb{R})^+$  is the identity embedding and  $\tilde{f}_\varrho: \mathbb{H} \rightarrow \mathbb{H}$  is the identity. Whenever we consider a specific modular embedding we will therefore tacitly assume that this is the case.

Sometimes we will need a stricter condition:

$$\varphi(\Delta) \text{ is contained in the image of } G(\mathbb{Q}) \rightarrow G^{\mathrm{ad}}(\mathbb{Q})^+. \quad (12)$$

Let  $\Delta \subset \mathrm{PGL}_2(\mathbb{R})^+$  be a lattice and let  $\tilde{\Delta}$  be its preimage in  $\mathrm{SL}_2(\mathbb{R})$ . The *trace field* of  $\Delta$  is the subfield

$$\mathbb{Q}(\mathrm{tr} \Delta) = \mathbb{Q}(\mathrm{tr} \tilde{\delta} \mid \tilde{\delta} \in \tilde{\Delta}) \subset \mathbb{R}$$

generated by all traces of elements of  $\tilde{\Delta}$ , and the *invariant trace field* of  $\Delta$  is the intersection of all trace fields of finite index subgroups of  $\Delta$ . If  $\Delta^{(2)}$  is the subgroup generated by all  $\delta^2$  for  $\delta \in \Delta$  (a normal subgroup of finite index), then the invariant trace field of  $\Delta$  equals  $\mathbb{Q}(\mathrm{tr} \Delta^{(2)})$ , see [30, Theorem 3.3.4].

**Proposition 4.2.** *Let  $(G, X)$  as above be associated with  $B/F$ . Let  $\Delta \subset \mathrm{PGL}_2(\mathbb{R})^+$  be a lattice such that there exists a modular embedding for  $\Delta$  with respect to  $\varrho \in \mathcal{P}$ .*

(i) *The invariant trace field of  $\Delta$  is contained in  $\varrho(F)$ . All traces  $\mathrm{tr} \tilde{\delta}$  for  $\tilde{\delta} \in \tilde{\Delta}$  are algebraic integers.*

(ii) *If (12) is true then also the trace field  $\mathbb{Q}(\mathrm{tr} \Delta)$  is contained in  $\varrho(F)$ .*

*Proof.* Let  $K$  be the invariant trace field and let  $\Delta' \subseteq \Delta$  be a subgroup of finite index with trace field  $K$ . Let  $A \subset \mathrm{M}_2(\mathbb{R})$  be the  $K$ -subalgebra generated by  $\tilde{\Delta}' \subset \mathrm{SL}_2(\mathbb{R}) \subset \mathrm{M}_2(\mathbb{R})$ ; this is a quaternion algebra over  $K$ .

By assumption the image of  $\Delta$  in  $G^{\mathrm{ad}}(\mathbb{Q})^+$  is contained in an arithmetic subgroup containing a group in  $\Sigma(G)$ . By further shrinking  $\Delta'$  we may assume that  $\varphi(\Delta') \subseteq \Gamma = \mathrm{ad} \tilde{F}$  for some  $\tilde{F} \in \tilde{\Sigma}(G)$ .

For (i), assume that  $\varphi_\varrho: \Delta \rightarrow \mathrm{PGL}_2(\mathbb{R})^+$  is the identity inclusion; then  $\tilde{\Delta}'$  is contained in the image of  $\tilde{F} \subset B^1 = G(\mathbb{Q})$  under the  $\varrho$ -th coordinate map in (5). But then  $K \subseteq \varrho(F)$  since the latter is the trace field of  $\tilde{F}$ , and the identity inclusion induces an isomorphism  $A \otimes_{K, \varrho^{-1}} F \cong B$ . The traces of elements in  $\tilde{\Delta}'$  are then of the form  $\varrho(x)$  for algebraic integers  $x \in \mathfrak{o}_F$ , hence the eigenvalues of any element in  $\tilde{\Delta}'$  are algebraic units. Since every element in  $\tilde{\Delta}$  has a finite nontrivial power in  $\tilde{\Delta}'$  also the eigenvalues of elements of  $\tilde{\Delta}$  are algebraic units, hence the traces are algebraic integers.

A very similar argument proves (ii).  $\square$

**Examples 4.3.** (i) If  $\Delta$  is an arithmetic group itself there is a tautological modular embedding. Recall that  $\Delta$  is arithmetic if and only if it is commensurable to a group of the form  $\mathcal{O}_0^1$  for some order  $\mathcal{O}_0$  in a quaternion algebra

$A_0/K_0$  which splits at precisely one archimedean place. Then there is a connected Shimura datum  $(G_0, X_0)$  with  $G_0(\mathbb{Q}) = A_0^1$  and a modular embedding  $(\tilde{f}, \varphi)$  with  $\tilde{f}: \mathbb{H} \rightarrow X_0$  a biholomorphism and  $\varphi: \Delta \rightarrow G_0^{\text{ad}}(\mathbb{Q})^+$  restricting to the canonical embedding on  $\Delta^{(2)}$ . Note that this can be extended to  $\Delta$  by the Skolem–Noether theorem which implies that  $\Delta$  is contained in the image of  $(A_0^\times)^+$  in  $\text{PGL}_2(\mathbb{R})^+$ , which is canonically isomorphic to  $G_0^{\text{ad}}(\mathbb{Q})^+$ .

We may, however, start with any totally real extension  $K/K_0$  and set  $A = A_0 \otimes_{K_0} K$ . From this we can easily construct a modular embedding for  $\Delta$  into  $(G, X)$  with  $G(\mathbb{Q}) = A^1$  and  $X \simeq \mathbb{H}^{[K:K_0]}$ ; each of its coordinates is (up to a change of coordinates) given by  $\tilde{f}$  and  $\varphi$  constructed before for  $(G_0, X_0)$ .

- (ii) For a specific example, let  $\Delta$  be the *Fricke group* for a rational prime  $p$ , which is the group generated by the images of

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(\mathbb{Z}) \mid c \equiv 0 \pmod{p} \right\} \quad \text{and} \quad \begin{pmatrix} 0 & -1 \\ p & 0 \end{pmatrix}$$

in  $\text{PGL}_2(\mathbb{Q})^+$ . Considered as a subgroup of  $\text{PGL}_2(\mathbb{R})^+$ , its invariant trace field is  $K_0 = \mathbb{Q}$  but its trace field is  $K = \mathbb{Q}(\sqrt{p})$  because the lifts of the additional matrix generator to  $\text{SL}_2(\mathbb{R})$  are

$$\pm \begin{pmatrix} 0 & -\frac{1}{\sqrt{p}} \\ \sqrt{p} & 0 \end{pmatrix}.$$

The associated quaternion algebra is  $B = \text{M}_2(K)$ , hence we can choose  $\Gamma$  to be a congruence subgroup of  $\text{SL}_2(K)$ , i.e. a *Hilbert–Blumenthal modular group* operating on  $\mathbb{H}^2$ . Now Galois conjugation operates trivially on the elements of  $\Delta$ , hence the diagonal map  $\mathbb{H} \rightarrow \mathbb{H}^2$  is a modular embedding for  $\Delta$ , and the image in the Hilbert–Blumenthal surface  $\Gamma \backslash \mathbb{H}^2$  is a *Hirzebruch–Zagier cycle*, cf. [23].

Note that the modular embedding into  $\text{PGL}_2(K)^+$  satisfies (12), but the tautological embedding into  $\text{PGL}_2(\mathbb{Q})^+$  does not.

- (iii) In a similar way we can construct a modular embedding where  $\Delta$  is cocompact but the arithmetic subgroups of  $G^{\text{ad}}(\mathbb{Q})^+$  are not. Let  $\Delta$  be a cocompact arithmetic group derived from a quaternion algebra  $A_0$  over  $K_0$ , and choose  $K$  to be a totally real extension of  $K_0$  that splits  $A_0$ . Then the construction in (ii) yields a modular embedding  $(\tilde{f}, \varphi)$  with  $\varphi: \Delta \rightarrow G^{\text{ad}}(\mathbb{Q})^+ = \text{PGL}_2(K)^+$ .

- (iv) Let  $p, q, r \in \mathbb{N} \cup \{\infty\}$  with

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} < 1,$$

where we set  $\frac{1}{\infty} = 0$ . Then there exists a lattice  $\Delta_{p,q,r} \subset \text{PGL}_2(\mathbb{R})^+$ , unique up to conjugacy, with presentation

$$\Delta_{p,q,r} = \langle x, y, z \mid x^p = y^q = z^r = xyz = 1 \rangle$$

where ‘ $x^\infty = 1$ ’ is interpreted as ‘ $x$  is a nontrivial parabolic element of  $\text{PGL}_2(\mathbb{R})^+$ ’, and similarly for  $y$  and  $z$ . This group is called the *triangle*

group of signature  $(p, q, r)$ . Its trace field is

$$K_{p,q,r} = \mathbb{Q}\left(\cos \frac{\pi}{p}, \cos \frac{\pi}{q}, \cos \frac{\pi}{r}\right)$$

and it generates a quaternion algebra  $A_{p,q,r}$  over this trace field. Then there is a connected quaternionic Shimura datum  $(G, X)$  with  $G(\mathbb{Q}) = A_{p,q,r}^1$  and a modular embedding consisting of  $\tilde{f}: \mathbb{H} \rightarrow X$  and  $\varphi: \Delta \rightarrow G^{\text{ad}}(\mathbb{Q})^+$ . This modular embedding is constructed in three different ways in [5], by Schwarz triangle functions, by hypergeometric differential equations and by period maps for special families of abelian varieties. It has the convenient property (12), as can be seen from the first of the three constructions in [5].

- (v) The uniformising groups of *Teichmüller curves*, i.e. algebraic curves which are immersed in the moduli space  $\mathcal{M}_g$  of smooth projective genus  $g$  curves in a totally geodesic way for the Teichmüller metric, admit modular embeddings. This was shown in [31, Theorem 10.1] for  $g = 2$  and in [38, Corollary 2.11] for the general case. These modular embeddings are studied in detail in [40]. The quaternion algebras in these modular embeddings are always of the form  $M_2(F)$  with  $\varrho(F)$  being equal to both the trace field and the invariant trace field. The equality of the latter is [25, Corollary 6.1.11]. This implies that they, too, have the property (12) that  $\varphi(\Delta)$  is contained in the image of  $G(\mathbb{Q}) = \text{SL}_2(F)$ .

A modular embedding  $\tilde{f}: \mathbb{H} \rightarrow X$  descends to a holomorphic map between the quotients  $f: \Delta \backslash \mathbb{H} \rightarrow \Gamma \backslash X$ . By a theorem of Borel [2, Theorem 3.10] this is induced by a regular map of normal algebraic varieties  $f: C \rightarrow S = \text{Sh}_\Gamma^0(G, X)$ . If  $f$  arises in this fashion we say it is *covered by a modular embedding*.

**4.2. Maps from curves to Shimura varieties.** Let  $(G, X)$  be a connected Shimura datum defined by a quaternion algebra  $B/F$  as before, let  $\Gamma \in \Sigma(G)$  and let  $S = \text{Sh}_\Gamma^0(G, X)$ . Let further  $C$  be a smooth complex algebraic curve and let  $f: C \rightarrow S$  be a morphism of algebraic varieties. From Liouville's Theorem we easily see that if the universal covering of  $C^{\text{an}}$  is  $\mathbb{C}$  or  $\mathbb{P}^1(\mathbb{C})$  then  $f$  is constant, so we assume from now on that  $C$  is hyperbolic.

After choosing appropriate basepoints we can lift  $f$  to a holomorphic map between the universal covers

$$\tilde{f}: \mathbb{H} \rightarrow X$$

which is equivariant for the induced map on fundamental groups

$$\tilde{\varphi} = f_*: \Delta = \pi_1(C^{\text{an}}) \rightarrow \Gamma \subset G(\mathbb{R}).$$

Using the classical description  $X \simeq \mathbb{H}^p$  and  $G(\mathbb{R}) \twoheadrightarrow \text{SL}_2(\mathbb{R})^p$  we can split these up into coordinate components

$$\tilde{f}_\varrho: \mathbb{H} \rightarrow \mathbb{H} \quad \text{and} \quad \tilde{\varphi}_\varrho: \pi_1(C^{\text{an}}) \rightarrow \text{SL}_2(\mathbb{R}).$$

We can also pull back the various bundles on  $S$  along  $f$ . Hence we obtain on  $C$ , for every  $\varrho \in \mathcal{P}$ , a real variation  $f^*\mathbb{W}_\varrho$  of elliptic curve type, a line bundle  $f^*\mathcal{L}_\varrho$  and a Higgs field

$$f^*\vartheta_\varrho: f^*\mathcal{L}_\varrho \rightarrow f^*\mathcal{L}_\varrho^{-1} \otimes f^*\Omega_S^1 \xrightarrow{(*)} f^*\mathcal{L}_\varrho^{-1} \otimes \omega_C, \quad (13)$$

where  $f^*\Omega_S^1 \rightarrow \Omega_C^1 = \omega_C$  in  $(*)$  is the canonical map. Note that since  $f$  is an algebraic map and the Higgs field  $\vartheta_\rho$  on  $S$  is algebraic, so is (13).

We first assume that  $C$  is a smooth projective curve.

**Theorem 4.4.** *Let  $C$  be a smooth projective curve of genus at least two, let  $S = \mathrm{Sh}_{\mathrm{ad}\tilde{\Gamma}}^0(G, X)$  with  $\tilde{\Gamma} \in \tilde{\Sigma}(G)$ , and let  $f: C \rightarrow S$  be a morphism of algebraic varieties. Fix some embedding  $\rho \in \mathcal{P}$ . With the notations of this section, the following are equivalent.*

- (i)  $\tilde{f}$  is a modular embedding with respect to  $\rho$ .
- (ii) The homomorphism  $\varphi_\rho: \pi_1(C^{\mathrm{an}}) \rightarrow \mathrm{SL}_2(\mathbb{R})$  is conjugate to a theta characteristic of  $\Delta = \pi_1(C^{\mathrm{an}})$ .
- (iii) The component map  $f_\rho: \mathbb{H} \rightarrow \mathbb{H}$  is a biholomorphism.
- (iv) The Higgs field  $f^*\vartheta_\rho$  is maximal.
- (v)  $f^*\mathcal{L}_\rho$  is a theta characteristic on  $C$ .
- (vi) The degree of the line bundle  $f^*\mathcal{L}_\rho$  is equal to  $-\frac{1}{2}\chi(C)$ .

If these conditions are satisfied, then  $\varphi_\rho \in \Theta^{\mathrm{uni}}(\Delta)$  in (ii) and  $f^*\mathcal{L}_\rho \in \Theta(C)$  in (v) correspond to each other under the bijection  $\Theta^{\mathrm{uni}}(\Delta) \leftrightarrow \Theta(C)$  from section 2.3.

*Proof.* This follows directly from Proposition 2.8.  $\square$

Finally we note that the line bundle  $\mathcal{M}_\rho = \mathcal{L}_\rho^{\otimes 2}$  is geometrically much easier to describe than  $\mathcal{L}_\rho$ : it is a summand in the decomposition

$$\Omega_S^1 = \bigoplus_{\rho \in \mathcal{P}} \mathcal{M}_\rho$$

coming from the product decomposition  $X \cong \mathbb{H}^p$ . Furthermore, it depends only on the group  $\Gamma \in \Sigma(G)$  instead of a choice of  $\tilde{\Gamma} \in \tilde{\Sigma}(G)$ .

**Corollary 4.5.** *Let  $C$  be a smooth projective curve of genus at least two, let  $S = \mathrm{Sh}_\Gamma^0(G, X)$  with  $\Gamma \in \Sigma(G)$  and let  $f: C \rightarrow S$  be a morphism of algebraic varieties. Fix some embedding  $\rho \in \mathcal{P}$ . Then  $f$  is covered by a modular embedding with respect to  $\rho$  if and only if  $f^*\mathcal{M}_\rho$  is isomorphic to  $\omega_C$  (as an algebraic, holomorphic or topological line bundle; these statements are all equivalent under these conditions).*

*Proof.* Assume first that  $f^*\mathcal{M}_{\rho, \Gamma} \simeq \omega_C$ , and let  $\Gamma' \subseteq \Gamma$  be a subgroup with  $\Gamma' = \mathrm{ad}\tilde{\Gamma}'$  for some  $\tilde{\Gamma}' \in \tilde{\Sigma}(G)$ . Let  $C' \rightarrow C$  be the finite unramified covering corresponding to the subgroup  $\varphi^{-1}(\Gamma') \subseteq \Delta$ , and let  $f': C' \rightarrow S' = \mathrm{Sh}_{\Gamma'}^0(G, X)$  be the obvious lift of  $f$ . Then by pullback  $f'^*\mathcal{M}_{\rho, \Gamma'} \simeq \omega_{C'}$ , hence  $f'^*\mathcal{L}_{\rho, \tilde{\Gamma}'}$  is a theta characteristic on  $C'$ . By Theorem 4.4  $f'$  is covered by a modular embedding, hence so is  $f$ .

Now assume that  $f$  is covered by a modular embedding, and choose  $\Gamma'$  as before, but in addition normal in  $\Gamma$ . Then we obtain an isomorphism of line bundles  $f'^*\mathcal{M}_{\rho, \Gamma} \rightarrow \omega_{C'}$  which is by construction invariant under the finite group  $\Gamma/\Gamma'$  acting on  $C'$ , hence it descends to an isomorphism  $f^*\mathcal{M}_{\rho, \Gamma} \rightarrow \omega_C$ .  $\square$

We now turn to the affine case.

**Proposition 4.6.** *If  $f: C \rightarrow S$  is covered by a modular embedding, where  $C$  is a smooth affine complex curve, then  $B \simeq \mathrm{M}_2(F)$ , hence  $S$  is a Hilbert–Blumenthal variety.*

Note that the reverse implication does not hold: there exist projective Shimura curves modularly embedded in non-projective Hilbert–Blumenthal varieties, cf. Example 4.3.(iii).

*Proof.* The preimage of the fundamental group of  $C^{\text{an}}$  in  $\text{SL}_2(\mathbb{R})$  contains nontrivial unipotent elements. Hence so does  $G(\mathbb{Q})$ ; but if  $1 \neq \gamma \in G(\mathbb{Q}) = B^1$  is unipotent, then  $\gamma - 1 \in B \setminus \{0\}$  is nilpotent, hence  $B$  cannot be a division algebra.  $\square$

We assume for the rest of this subsection that  $B = M_2(F)$ . Recall that then  $\mathbb{W}_\varrho$  can be described as a direct summand, defined by an eigenvalue equation, of the variation  $R^1p_*\mathbb{R}$ , where  $p: A \rightarrow S$  is the universal family, and the associated vector bundle with its Hodge filtration and its connection can be constructed in a purely algebraic way from the universal family.

**Lemma 4.7.** *After possibly replacing  $\Gamma$  by a suitable subgroup in  $\Sigma(G)$  and correspondingly  $C$  by a finite unramified covering we may assume that  $f^*\mathbb{W}_\varrho$  has unipotent monodromy around the cusps of  $C$ .*

*Proof.* Let  $p_C: A_C \rightarrow C$  be the pullback of the universal family  $p: A \rightarrow S$  along  $f$ . Then by (10) the variation  $f^*\mathbb{W}_\varrho$  is a direct summand of the variation  $R^1p_{C,*}\mathbb{R}$  on  $C$ , and such variations have quasi-unipotent monodromy around the cusps by [7, III.§2, Théorème 2.3].  $\square$

**Theorem 4.8.** *Let  $S$  and  $C = \overline{C} \setminus D$  as before, and let  $f: C \rightarrow S$  be a regular map corresponding to an algebraic family of abelian varieties  $p_C: A_C \rightarrow C$  with an inclusion  $F \hookrightarrow \text{End}(A_C/C) \otimes \mathbb{Q}$ . Denote  $\varrho$ -isotypical components of  $F$ -modules by subscripts  $\varrho$ . Then, after possibly passing to finite unramified coverings, the Gauss–Manin connection  $\nabla_\varrho^{\text{GM}}$  on the vector bundle  $(\mathcal{H}_{\text{dR}}^1(A_C/C))_\varrho$  on  $C$  has regular singularities with unipotent monodromy around the cusps of  $C$ . Let  $\eta_{\text{ext}}: \mathcal{E}_{\text{ext}} \rightarrow \mathcal{E}_{\text{ext}}^{-1} \otimes \Omega_{\overline{C}}(\log D)$  be the logarithmic Higgs bundle corresponding to the Deligne extension of  $(\mathcal{H}_{\text{dR}}^1(A_C/C))_\varrho$ .*

*Then the following are equivalent:*

- (i)  *$f$  is covered by a modular embedding for  $\varrho$ .*
- (ii)  *$\eta_{\text{ext}}$  is an isomorphism.*
- (iii)  *$\mathcal{E}_{\text{ext}}$  is a logarithmic theta characteristic on  $C$ .*
- (iv) *The degree of  $\mathcal{E}_{\text{ext}}$  is  $-\frac{1}{2}\chi(C)$ .*

*Proof.* Since the logarithmic Higgs field  $\eta_{\text{ext}}$  is isomorphic to that associated with the variation  $f^*\mathbb{W}_\varrho$ , conditions (ii) and (iii) can be reformulated as statements about that variation. The  $\varrho$ -th component  $\tilde{f}_\varrho: \mathbb{H} \rightarrow \mathbb{H}$  of the universal covering map  $\tilde{f}$  is the period map for this variation. The theorem follows from Proposition 2.8.  $\square$

**Remark 4.9.** It is desirable to compare Theorems 4.4 and 4.8. Note that in the setup of Theorem 4.8 the line bundle  $\mathcal{E} = \mathcal{E}_{\text{ext}}|_C = p_*(\Omega_{A_C/C}^1)_\varrho$  is isomorphic as an algebraic line bundle to  $f^*\mathcal{L}_\varrho$ , with the Higgs fields  $\eta = \eta_{\text{ext}}|_C$  and  $f^*\vartheta_\varrho$  corresponding to each other. Once we know that  $f$  is covered by a modular embedding, there is a unique extension of  $(f^*\mathcal{L}_\varrho, f^*\vartheta_\varrho)$  to a logarithmic Higgs bundle on  $\overline{C}$  with maximal Higgs field, but for an arbitrary regular map  $f$  the situation is more complicated and there is no obvious way to choose an extension of  $(f^*\mathcal{L}_\varrho, f^*\vartheta_\varrho)$  to a logarithmic Higgs bundle on  $\overline{C}$ . It seems that we need the subtler information

encoded in the relative de Rham cohomology to decide how to extend the Higgs bundle to the cusps. Alternatively one might study a suitable extension  $\tilde{f}: \overline{C} \rightarrow \overline{S}$  for some compactification  $\overline{S}$  of  $S$ , corresponding to a family of semi-abelian varieties over  $\overline{C}$ , and try to obtain  $\eta_{\text{ext}}$  as a Kodaira–Spencer map for  $\tilde{f}$ . This, however, seems technically much more difficult and is not necessary for our purposes.

**4.3. Adelic theory.** Let  $\Delta \subset \text{PGL}_2(\mathbb{R})^+$  be a semi-arithmetic lattice, let  $\Gamma \subset G^{\text{ad}}(\mathbb{Q})^+$  be an arithmetic subgroup containing the image of a congruence subgroup of  $G(\mathbb{Q})$ , and let  $\varphi: \Delta \rightarrow \Gamma$  and  $\tilde{f}: \mathbb{H} \rightarrow X$  constitute a modular embedding with respect to  $\varrho: F \rightarrow K \subset \mathbb{R}$ .

**Definition 4.10.** Let  $\Delta' \subseteq \Delta$  be a subgroup. Then  $\Delta'$  is a congruence subgroup of  $\Delta$  with respect to  $\varphi$  if it contains a subgroup of the form  $\varphi^{-1}(\Gamma')$  with  $\Gamma' \in \Sigma(G)$ .

Recall that  $\mathcal{O} = \mathfrak{o}_K \langle \Delta \rangle$  is an order in a quaternion algebra  $A/K$ . For every ideal  $\mathfrak{n}$  of  $K$  we define the *principal congruence subgroup*  $\Delta(\mathfrak{n})$  as

$$\Delta(\mathfrak{n}) = \{\delta \in \Delta \mid \delta - 1 \in \mathfrak{n}\mathcal{O}\}.$$

This is clearly a congruence subgroup.

**Proposition 4.11.** Assume that  $\varphi$  satisfies (12). Then a subgroup of  $\Delta$  is a congruence subgroup with respect to  $\varphi$  if and only if it contains a subgroup of the form  $\Delta(\mathfrak{n})$ .

*Proof.* Easy. □

Therefore for modular embeddings satisfying (12) the notion of congruence subgroup is intrinsic to  $\Delta$ . In general it can be slightly more complicated. In any case the congruence subgroups define the neighbourhood basis of the identity for a topological group structure on  $\Delta$ , and the completion  $\hat{\Delta}$ , called the *congruence completion*, is canonically isomorphic to

$$\hat{\Delta} \cong \varprojlim_{\Delta'} \Delta/\Delta',$$

$\Delta'$  running over all normal congruence subgroups of  $\Delta$ . This is a profinite group. Similarly, the congruence completion of  $\Gamma$  is a profinite group  $\hat{\Gamma}$  which is canonically isomorphic to the topological closure of  $\Gamma$  in  $\mathcal{A}(G)$ , an open compact subgroup of the latter. The homomorphism  $\varphi: \Delta \rightarrow \Gamma$  has a unique continuous extension  $\hat{\varphi}: \hat{\Delta} \rightarrow \mathcal{A}(G)$ . It is injective, and it induces a homeomorphism from  $\hat{\Delta}$  to a closed subgroup of  $\hat{\Gamma}$ . Under the additional condition that the trace field of  $\Delta$  is the same as the invariant trace field this closed subgroup is in fact open by superstrong approximation, cf. [26, Proposition 4.4].

For a subgroup  $\Gamma \in \Sigma(G)$  let  $C_{\varphi^{-1}(\Gamma)}$  be the complex curve with  $C_{\varphi^{-1}(\Gamma)}(\mathbb{C}) = \varphi^{-1}(\Gamma) \backslash \mathbb{H}$ . Then  $\tilde{f}$  descends to a morphism of algebraic varieties

$$f_{\Gamma}: C_{\varphi^{-1}(\Gamma)} \rightarrow \text{Sh}_{\Gamma}^0(G, X).$$

Passing to limits in the category of  $\mathbb{C}$ -schemes, we obtain  $\mathbb{C}$ -schemes of infinite type

$$\hat{C} \stackrel{\text{def}}{=} \varprojlim_{\Gamma} C_{\varphi^{-1}(\Gamma)} \quad \text{and} \quad \text{Sh}^0(G, X) \cong \varprojlim_{\Gamma} \text{Sh}_{\Gamma}^0(G, X).$$

By construction,  $\hat{\Delta}$  acts continuously by  $\mathbb{C}$ -scheme isomorphisms on  $\hat{C}$ , and  $\mathcal{A}(G)$  acts similarly on  $\mathrm{Sh}^0(G, X)$  (as we have already seen).

We summarise these considerations:

**Proposition 4.12.** *A modular embedding as above gives rise to a morphism of schemes of infinite type over  $\mathbb{C}$ ,*

$$\hat{f}: \hat{C} \rightarrow \mathrm{Sh}^0(G, X),$$

*which is equivariant for the embedding  $\hat{\varphi}: \hat{\Delta} \rightarrow \mathcal{A}(G)$ .*

*For a congruence subgroup  $\Delta' \subseteq \Delta$  with closure  $\hat{\Delta}'$  in  $\hat{\Delta}$  the corresponding curve  $C'$  can be retrieved as  $C' \cong \hat{\Delta}' \backslash \hat{C}$ , and similarly for  $\mathrm{Sh}^0(G, X)$ .*

## 5. GALOIS CONJUGATION

In this section we study how abstract field automorphisms of  $\mathbb{C}$  act on Shimura varieties, automorphic Higgs bundles and modular embeddings. In the end we draw some consequences for dessins d'enfants.

The action of field automorphisms of  $\mathbb{C}$  on Shimura varieties and automorphic bundles can be described in elementary terms for Shimura varieties of type PEL, but for more general Shimura varieties such as the ones we are considering a more elaborate approach is needed. This approach constructs, for each connected Shimura datum  $(G, X)$ , each field automorphism  $\tau \in \mathrm{Aut} \mathbb{C}$  and each special point  $x \in X$ , a new Shimura datum  $({}^{\tau, x}G, {}^{\tau, x}X)$  which is obtained from the original one by twisting the groups involved by certain torsors that can be described by Galois cohomology. We give a review of this theory adapted to our purposes in sections 5.1 to 5.3.

**5.1. The Serre, Taniyama and Galois groups.** Fundamental to the general theory of Galois conjugation on Shimura varieties is a short exact sequence

$$1 \rightarrow \mathfrak{S} \rightarrow \mathfrak{T} \rightarrow \mathfrak{G} \rightarrow 1 \tag{14}$$

of pro-algebraic groups (i.e., projective limits of linear algebraic groups) over  $\mathbb{Q}$ , equipped with certain splitting data over  $\mathbb{C}$  and  $\mathbb{Q}_p$ . We now review the construction of these three groups, the sequence (14) and various auxiliary structures.

*Elementary construction of the groups.* The group  $\mathfrak{G}$  is simply the *absolute Galois group*  $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ , interpreted as a constant profinite group scheme over  $\mathbb{Q}$ . That is,

$$\mathfrak{G} = \varprojlim_L (\mathrm{Gal}(L/\mathbb{Q}))_{\mathbb{Q}},$$

where  $L$  runs through all finite Galois extensions of  $\mathbb{Q}$  contained in  $\overline{\mathbb{Q}}$ , and  $G_{\mathbb{Q}}$  denotes the constant group scheme  $\mathrm{Spec} \mathbb{Q}^G$  for any finite group  $G$ .

The group  $\mathfrak{S}$  is the *Serre group* (also known as Serre torus); it can also be constructed as a projective limit

$$\mathfrak{S} = \varprojlim_L \mathfrak{S}^L. \tag{15}$$

The facts mentioned below are proved in the opening section of [37].

The group  $\mathfrak{S}^L$  is the quotient of the torus  $\mathbb{T}^L = \mathrm{Res}_{L/\mathbb{Q}} \mathbb{G}_m$  by the Zariski closure of any sufficiently small finite index subgroup of  $\mathfrak{o}_L^\times \subset L^\times = \mathbb{T}^L(\mathbb{Q})$ . The character group of  $\mathbb{T}^L$  is the free  $\mathbb{Z}$ -module spanned by  $[\varrho]$ , where  $\varrho$  runs through all field

embeddings  $\varrho: L \rightarrow \overline{\mathbb{Q}}$ , with Galois action  $\sigma \cdot [\varrho] = [\sigma \circ \varrho]$ . Hence the character group of  $\mathfrak{S}^L$  is a Galois-stable subgroup of this group; an easy argument involving Minkowski's Geometry of Numbers shows that  $X^*(\mathfrak{S}^L)$  consists precisely of the characters  $\chi \in X^*(\mathbb{T}^L)$  with

$$(\sigma - 1)(\iota + 1)\chi = (\iota + 1)(\sigma - 1)\chi = 0 \text{ for all } \sigma \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}). \quad (16)$$

This condition can be reformulated as follows:

$$X^*(\mathfrak{S}^L) = \left\{ \sum_{\varrho: L \rightarrow \overline{\mathbb{Q}}} a_{\varrho}[\varrho] \mid -a_{\varrho} - a_{\iota\varrho} \text{ independent of } \varrho \right\}. \quad (17)$$

There is a natural homomorphism  $h^L: \mathbb{S} \rightarrow \mathfrak{S}_{\mathbb{R}}^L$  given on character groups by

$$X^*(h^L): X^*(\mathfrak{S}^L) \rightarrow X^*(\mathbb{S}) = \bigoplus_{\varrho': \mathbb{R} \rightarrow \mathbb{C}} \mathbb{Z}[\varrho'], \quad \sum_{\varrho} a_{\varrho}[\varrho] \mapsto a_{\text{id}}[\text{id}] + a_{\iota}[\iota].$$

Composition with the ‘weight’ and ‘Hodge filtration’ homomorphisms yields homomorphisms  $w^L: \mathbb{G}_m \rightarrow \mathfrak{S}^L$  (nota bene: defined over  $\mathbb{Q}$ ) and  $\mu^L: \mathbb{G}_{m,\mathbb{C}} \rightarrow \mathfrak{S}_{\mathbb{C}}^L$  given on characters by

$$X^*(w^L): X^*(\mathfrak{S}^L) \rightarrow \mathbb{Z}, \quad \sum_{\varrho} a_{\varrho}[\varrho] \mapsto -a_{\text{id}} - a_{\iota} = \text{the constant from (17),}$$

and

$$X^*(\mu^L): X^*(\mathfrak{S}^L) \rightarrow \mathbb{Z}, \quad \sum_{\varrho} a_{\varrho}[\varrho] \mapsto a_{\text{id}}.$$

**Definition 5.1.** *Let  $T$  be a  $\mathbb{Q}$ -torus which splits over  $L$ , and let  $\mu: \mathbb{G}_{m,\mathbb{C}} \rightarrow T_{\mathbb{C}}$  be a cocharacter. Then  $\mu$  satisfies the Serre condition if*

$$(\sigma - 1)(\iota + 1)\mu = (\iota + 1)(\sigma - 1)\mu = 0 \quad \text{for all } \sigma \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$$

(compare (16)).

A simple calculation shows that the pair  $(\mathfrak{S}^L, \mu^L)$  is universal for tori split over  $L$  with a cocharacter satisfying the Serre condition:

**Lemma 5.2.** *Let  $T$  be a  $\mathbb{Q}$ -torus which splits over  $L$ , and let  $\mu \in X_*(T)$  be a cocharacter satisfying the Serre condition. Then there exists a unique morphism of tori  $\alpha: \mathfrak{S}^L \rightarrow T$  (defined over  $\mathbb{Q}$ ) such that  $\mu = \alpha_{\mathbb{C}} \circ \mu^L$ .*

For each finite extension  $L \subseteq L' \subset \overline{\mathbb{Q}}$  the norm map  $N_{L'/L}$  restricts to a homomorphism  $\mathfrak{S}^{L'} \rightarrow \mathfrak{S}^L$ , and the limit (15) is to be understood with respect to these homomorphisms. Furthermore we can restrict ourselves to CM-fields  $L$  because  $\mathfrak{S}^L \cong \mathfrak{S}^{L'}$  if  $L$  is the union of all CM- or totally real subfields of  $L'$ . Hence  $\mathfrak{S}$  is the pro-torus whose character group  $X^*(\mathfrak{S})$  consists of all functions  $n: \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \rightarrow \mathbb{Z}$  which factor through some  $\text{Gal}(L/\mathbb{Q})$  with  $L/\mathbb{Q}$  Galois and CM and for which  $-n(\sigma) - n(\iota\sigma)$  is independent of  $\sigma$ . Since the homomorphisms  $h^L$ ,  $w^L$  and  $\mu^L$  are compatible with this limiting process, we obtain homomorphisms

$$h_{\text{can}} = \lim_L h^L: \mathbb{S} \rightarrow \mathfrak{S}_{\mathbb{R}}, \quad w_{\text{can}}: \mathbb{G}_m \rightarrow \mathfrak{S}, \quad \mu_{\text{can}}: \mathbb{G}_{m,\mathbb{C}} \rightarrow \mathfrak{S}_{\mathbb{C}}.$$

From Lemma 5.2 we deduce that  $(\mathfrak{S}, \mu_{\text{can}})$  is universal among pairs  $(T, \mu)$  where  $T$  is a torus over  $\mathbb{Q}$  and  $\mu \in X_*(T)$  satisfies the Serre condition.

The group  $\mathfrak{T}$  is called the *Taniyama group*. It can also be constructed as a limit object. In [28, section 5] Langlands constructs, for each finite Galois extension  $L/\mathbb{Q}$  contained in  $\overline{\mathbb{Q}}$ , an exact sequence of groups

$$1 \rightarrow \mathfrak{S}^L \rightarrow \mathfrak{T}^L \rightarrow \mathfrak{G}^L \rightarrow 1, \quad (18)$$

where  $\mathfrak{G}^L$  is the constant group  $\mathbb{Q}$ -scheme defined by  $\text{Gal}(L^{\text{ab}}/\mathbb{Q})$ , and then shows that (18) admits a limit as  $L$  varies.

*Tannakian theory and motives.* The sequence (14) also admits another description in terms of Tannakian categories. This is much less amenable to explicit calculations, but it is conceptually very elegant and explains the appearance of this sequence when field automorphisms are applied to Shimura varieties. For the formalism of Tannakian categories and fibre functors see [13]. The theory to be summarised below is beautifully explained in [34].

Let  $\mathbf{CM}_{\mathbb{C}}$  be the Tannakian subcategory of the category of pure Hodge  $\mathbb{Q}$ -structures generated by all  $H_1(A^{\text{an}}, \mathbb{Q})$ , where  $A$  is an abelian variety with complex multiplication over  $\mathbb{C}$ , and the Tate Hodge structure  $\mathbb{Q}(1)$ . Then the Tannakian fundamental group of  $\mathbf{CM}_{\mathbb{C}}$  for the fibre functor given by the underlying  $\mathbb{Q}$ -vector spaces is canonically isomorphic to  $\mathfrak{S}$ , see [12] or [34, Prop. 4.5].

We let  $\mathbf{Art}_{\mathbb{Q}}$  be the category of Artin motives, i.e. of finite-dimensional continuous  $\mathbb{Q}$ -representations of  $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ . Then the Tannakian fundamental group of  $\mathbf{Art}_{\mathbb{Q}}$  with respect to the tautological fibre functor is canonically isomorphic to  $\mathfrak{S}$ .

Finally we let  $\mathbf{CM}_{\mathbb{Q}}$  be the Tannakian category of motives over  $\mathbb{Q}$  generated by the Tate motive  $\mathbb{Q}(1) = H_2(\mathbb{P}^1)$ , the motives  $H_1(A)$  for all abelian varieties  $A$  over  $\mathbb{Q}$  with potential complex multiplication, and the Artin motives (which can be viewed as subobjects of the motives  $H_0(V)$ , where  $V$  runs over all finite  $\mathbb{Q}$ -varieties). All these motives can be constructed unconditionally using the theory of absolute Hodge cycles developed in [14]. Then we let  $\mathfrak{T}$  be the fundamental group of  $\mathbf{CM}_{\mathbb{Q}}$  for the fibre functor given by the Betti realisation.

The inclusion  $\mathbf{Art}_{\mathbb{Q}} \hookrightarrow \mathbf{CM}_{\mathbb{Q}}$  and the base change  $\mathbf{CM}_{\mathbb{Q}} \rightarrow \mathbf{CM}_{\mathbb{C}}$  give rise to the homomorphisms in (14).

*Splittings.* For every finite prime  $\ell$  we can identify  $\mathfrak{S}(\mathbb{Q}_{\ell})$  with  $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ , hence the  $\mathbb{Q}_{\ell}$ -valued points of (14) form an exact sequence

$$1 \rightarrow \mathfrak{S}(\mathbb{Q}_{\ell}) \rightarrow \mathfrak{T}(\mathbb{Q}_{\ell}) \rightarrow \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q});$$

the last map is in fact also surjective, and there is even a canonical splitting

$$\text{sp}_{\ell}: \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \rightarrow \mathfrak{T}(\mathbb{Q}_{\ell}).$$

It can easily be defined using the Tannakian formalism:

Let  $H_B$  be the canonical ‘Betti’ fibre functor  $\mathbf{CM}_{\mathbb{Q}} \rightarrow \mathbf{Vect}_{\mathbb{Q}}$ . Then  $H_{\ell} = H_B \otimes_{\mathbb{Q}} \mathbb{Q}_{\ell}$  is endowed with a Galois representation (arising from the Galois action on Tate modules of abelian varieties). Hence each  $\tau \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$  gives rise to an automorphism of  $H_{\ell}$ , hence to an element of  $\mathfrak{T}(\mathbb{Q}_{\ell})$ , and this is by definition  $\text{sp}_{\ell}(\tau)$ . Note that  $\text{sp}_{\ell}$  does not come from a homomorphism of pro-algebraic groups  $\mathfrak{S}_{\mathbb{Q}_{\ell}} \rightarrow \mathfrak{T}_{\mathbb{Q}_{\ell}}$ , see [12, p. 261].

The product  $\text{sp}: \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \rightarrow \mathfrak{T}(\mathbb{A}^f)$  of the  $\text{sp}_{\ell}$  is likewise a splitting of the sequence of groups of  $\mathbb{A}^f$ -valued points obtained from (14).

**5.2. Torsors and Galois cohomology.** We recall some basic constructions in Galois cohomology, in both explicit and cohomological descriptions. For details and proofs, see the classical source [45, Chap. III, §1] or, in greater detail, the recent [16].

Recall that for an algebraic group  $G$  over a field  $k$ , a (*right*) *torsor* or *principal homogeneous space* is a  $k$ -variety  $P$  together with a right action of  $G$  on  $P$  such that over an algebraic closure  $\bar{k}$  there exists an isomorphism  $P_{\bar{k}} \rightarrow G_{\bar{k}}$  equivariant for  $G_{\bar{k}}$ , which operates on itself by right translations. A torsor is *trivial* if such an isomorphism already exists over  $k$ ; this is the case if and only if  $P(k)$  is nonempty.

Let  $X$  be a  $k$ -variety on which  $G$  acts from the left. Then the *contracted product* or *twist of  $X$  by  $P$* , denoted by  $P \times^G X$ , is the scheme-theoretic quotient  $(P \times_k X)/G$  (provided that it exists), where  $g \in G(\bar{k})$  acts by  $(p, x) \mapsto (pg^{-1}, gx)$ . Two special cases shall be of interest:

- (i) If  $f: G \rightarrow H$  is a homomorphism of algebraic groups, then  $P \times^G H$  (where  $H$  is acted upon by  $G$  by left multiplication via  $f$ ) is in a natural way a right  $H$ -torsor, denoted by  $f_*P$ .
- (ii) If  $X$  has some extra structure respected by the  $G$ -action (such as a group structure, or a  $k$ -algebra structure on a variety isomorphic to  $\mathbb{A}_k^n$ ), then  $P \times^G X$  has the same structure.

The  $G$ -torsors up to isomorphism are classified by the non-abelian Galois cohomology set  $H^1(k, G)$ . If  $G$  is the automorphism group of some ‘variety with additional structures’  $X$ , then the same cohomology set classifies forms of  $X$ , i.e. varieties with additional structures  $X'$  with  $X'_k \simeq X_k$ . This correspondence is realised by sending a torsor  $P$  to  $P \times^G X$ , and pushforward of torsors as in (i) corresponds to pushforward of cohomology classes  $f_*: H^1(k, G) \rightarrow H^1(k, H)$ .

Returning to the exact sequence (14), the Tannakian formalism gives rise to a simple construction of a  $\mathfrak{T}$ -torsor, the *period torsor*  $\mathfrak{P} = \underline{\text{Isom}}^{\otimes}(\mathbf{H}_B, \mathbf{H}_{\text{dR}})$ . Here  $\mathbf{H}_B$  is the fibre functor on  $\mathbf{CM}_{\mathbb{Q}}$  defined by the Betti realisation, and  $\mathbf{H}_{\text{dR}}$  is defined by the de Rham realisation. The comparison isomorphism  $\mathbf{H}_B \otimes_{\mathbb{Q}} \mathbb{C} \cong \mathbf{H}_{\text{dR}} \otimes_{\mathbb{Q}} \mathbb{C}$  defines an element  $c \in \mathfrak{P}(\mathbb{C})$ , and for each  $\tau \in \text{Aut } \mathbb{C}$  we let  $z_{\infty}(\tau) \in \mathfrak{T}(\mathbb{C})$  be the unique element with

$$c \cdot z_{\infty}(\tau) = \tau(c) \in \mathfrak{P}(\mathbb{C}).$$

The map  $z_{\infty}: \text{Aut } \mathbb{C} \rightarrow \mathfrak{T}(\mathbb{C})$  is a *crossed* homomorphism, and it does not factor through  $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ .

For  $\tau \in \text{Aut } \mathbb{C}$  the preimage of  $\{\tau|_{\overline{\mathbb{Q}}}\} \subset \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) = \mathfrak{G}(\mathbb{Q})$  under  $\mathfrak{T} \rightarrow \mathfrak{G}$  is a  $\mathbb{Q}$ -subscheme  ${}^{\tau}\mathfrak{S} \subset \mathfrak{T}$  which is by construction a right  $\mathfrak{S}$ -torsor. In fact, the point  $z_{\infty}(\tau)$  lies in  ${}^{\tau}\mathfrak{S}(\mathbb{C}) \subset \mathfrak{T}(\mathbb{C})$  (see [33, Prop. 1.1]) and therefore provides a distinguished trivialisation of  ${}^{\tau}\mathfrak{S}$  over  $\mathbb{C}$ . For further reference let  $s_{\tau} \in H^1(\mathbb{Q}, \mathfrak{S})$  be the cohomology class corresponding to  ${}^{\tau}\mathfrak{S}$ .

**5.3. Galois conjugates of Shimura varieties.** We recall the action of  $\tau \in \text{Aut } \mathbb{C}$  on the connected Shimura variety  $\text{Sh}^0(G, X)$ . The description of this action depends on a special point  $x \in X$  that is for now considered arbitrary and will later be specified conveniently.

*Special points.* Let  $E \subset B$  be an  $F$ -subalgebra which is a CM field extension of  $F$ . Then let  $\tilde{T} \subset G$  be the maximal torus with

$$\tilde{T}(\mathbb{Q}) = \{e \in E^\times \mid N_{E/F}(e) = 1\} \subset B^1 = G(\mathbb{Q}),$$

and let

$$T = \tilde{T}/\mathcal{Z}(G) \cong \mathbb{T}^E/\mathbb{T}^F$$

be its image in  $G^{\text{ad}} = G/\mathcal{Z}(G)$ ; that is,

$$T(\mathbb{Q}) = E^\times/F^\times \subset B^\times/F^\times = G(\mathbb{Q}).$$

By an *incomplete CM-type* on  $E$  we mean a subset  $\Phi \subset \text{Hom}_{\mathbb{Q}}(E, \mathbb{C})$  such that  $\Phi \cap \iota\Phi = \emptyset$  and

$$\{\varphi|_F^{\mathbb{R}} \mid \varphi \in \Phi\} = \mathcal{P} \subseteq \text{Hom}_{\mathbb{Q}}(F, \mathbb{R}).$$

There is a unique fixed point  $x$  of  $T(\mathbb{R})$  on  $X$ . The tangent space  $\text{Tgt}_x X$  is a complex vector space of dimension  $|\mathcal{P}|$  with a linear action of  $E^\times$  via its quotient  $T(\mathbb{Q})$ . The set  $\Phi$  of embeddings  $E \rightarrow \mathbb{C}$  that appear as a character in the  $E^\times$ -representation on  $\text{Tgt}_x X$  is an incomplete CM-type on  $E$ , and as an  $E^\times$ -module  $\text{Tgt}_x X$  is isomorphic to

$$\bigoplus_{\varphi \in \Phi} \mathbb{C}_{\varphi}.$$

Since  $T$  is constructed as a quotient of  $\mathbb{T}^E = \text{Res}_{E/\mathbb{Q}} \mathbb{G}_m$ , its character group becomes identified with a  $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ -submodule of

$$X^*(\mathbb{T}^E) \cong \bigoplus_{\sigma: E \rightarrow \mathbb{C}} \mathbb{Z} \cdot [\sigma],$$

namely with

$$X^*(T) \cong \left\{ \sum_{\sigma: E \rightarrow \mathbb{C}} a_{\sigma} [\sigma] \mid a_{\sigma} \in \mathbb{Z}, a_{\sigma} + a_{\iota\sigma} = 0 \text{ for all } \sigma \right\}.$$

Let  $\tilde{h}_x: \mathbb{S} \rightarrow (\mathbb{T}^E)_{\mathbb{R}}$  be the homomorphism of real algebraic tori given on character groups by

$$\bigoplus_{\sigma: E \rightarrow \mathbb{C}} \mathbb{Z} \cdot [\sigma] \rightarrow X^*(\mathbb{S}) = \mathbb{Z} \cdot [\text{id}] \oplus \mathbb{Z} \cdot [\iota], \quad [\sigma] \mapsto \begin{cases} [\text{id}] - [\iota] & \text{if } \sigma \in \Phi, \\ [\iota] - [\text{id}] & \text{if } \sigma \in \iota\Phi, \\ 0 & \text{if } \sigma \notin \Phi \cup \iota\Phi, \end{cases}$$

and let  $h_x: \mathbb{S} \rightarrow T_{\mathbb{R}}$  be its composition with  $\mathbb{T}^E \rightarrow T$ . An easy calculation shows that the composition  $h_x: \mathbb{S} \rightarrow T_{\mathbb{R}} \subset G_{\mathbb{R}}$  is precisely the homomorphism corresponding to the point  $x \in X$ , when  $X$  is interpreted as a  $G^{\text{ad}}(\mathbb{R})^+$ -conjugacy class of homomorphisms  $\mathbb{S} \rightarrow G_{\mathbb{R}}^{\text{ad}}$ .

The corresponding cocharacter  $\tilde{\mu}_x: \mathbb{G}_{m, \mathbb{C}} \rightarrow (\mathbb{T}^E)_{\mathbb{C}}$  which defines the Hodge filtration is given by

$$X^*(\mathbb{T}^E) \rightarrow X^*(\mathbb{G}_m) = \mathbb{Z}, \quad [\sigma] \mapsto \begin{cases} 1 & \text{if } \sigma \in \Phi, \\ -1 & \text{if } \sigma \in \iota\Phi, \\ 0 & \text{if } \sigma \notin \Phi \cup \iota\Phi; \end{cases} \quad (19)$$

this uniquely determines  $\mu_x: \mathbb{G}_{m, \mathbb{C}} \rightarrow T_{\mathbb{C}}$ . The cocharacter  $\mu_x$  (but not necessarily  $\tilde{\mu}_x$ , see Remark 5.6.(i)) satisfies the Serre condition and hence factors through the

Serre torus. That is, there is a unique homomorphism  $\alpha_x: \mathfrak{S} \rightarrow T$  defined over  $\mathbb{Q}$  such that the diagram

$$\begin{array}{ccc} \mathfrak{S} & \xrightarrow{h_x} & T_{\mathbb{R}} \\ & \searrow h_{\text{can}} & \nearrow (\alpha_x)_{\mathbb{R}} \\ & \mathfrak{S}_{\mathbb{R}} & \end{array}$$

commutes.

*The twisted group.* The pushforward of the  $\mathfrak{S}$ -torsor  ${}^{\tau}\mathfrak{S}$  along the composition

$$\alpha_x: \mathfrak{S} \rightarrow T \subset G^{\text{ad}}$$

defines a cohomology class  $g_{\tau} = \alpha_{x,*}(s_{\tau}) \in H^1(\mathbb{Q}, G^{\text{ad}})$  and, by interpreting  $G^{\text{ad}}$  as the inner automorphism group of  $G$ , an inner form  ${}^{\tau,x}G = {}^{\tau}\mathfrak{S} \times^{\mathfrak{S}} G$  of  $G$ . Similarly we obtain the inner form  ${}^{\tau,x}G^{\text{ad}} = {}^{\tau}\mathfrak{S} \times^{\mathfrak{S}} G^{\text{ad}}$  of  $G^{\text{ad}}$ .

But  $G^{\text{ad}}$  is also the group of  $F$ -algebra automorphisms of  $B$ , hence twisting by  $\alpha_{x,*}({}^{\tau}\mathfrak{S})$  defines an  $F$ -algebra  ${}^{\tau,x}B = B \times^{{}^{\tau}\mathfrak{S}} \mathfrak{S}$ . It is a form of  $B$ , hence a quaternion algebra over  $F$ ; the group  ${}^{\tau,x}G^{\text{ad}}(\mathbb{Q})$  can be canonically identified with  ${}^{\tau,x}B^{\times}/F^{\times}$ , and similarly for  ${}^{\tau,x}G$ . The point  $\text{sp}(\tau|_{\overline{\mathbb{Q}}}) \in {}^{\tau}\mathfrak{S}(\mathbb{A}^f)$  trivialises the torsor  ${}^{\tau}\mathfrak{S}$  over  $\mathbb{A}^f$ , hence it defines a canonical isomorphism of algebras over  $F \otimes_{\mathbb{Q}} \mathbb{A}^f \cong \mathbb{A}_F^f$ :

$$\eta: B \otimes_F \mathbb{A}_F^f \rightarrow {}^{\tau,x}B \otimes_F \mathbb{A}_F^f.$$

It can be split into local components

$$\eta_v: B \otimes_F F_v \rightarrow {}^{\tau,x}B \otimes_F F_v \tag{20}$$

for every finite prime  $v$  of  $F$ . We will also denote the corresponding isomorphisms

$$G_{\mathbb{A}^f} \rightarrow {}^{\tau,x}G_{\mathbb{A}^f} \quad \text{and} \quad G_{\mathbb{A}^f}^{\text{ad}} \rightarrow {}^{\tau,x}G_{\mathbb{A}^f}^{\text{ad}} \tag{21}$$

by  $\eta$ . There is also a canonical isomorphism of topological groups

$$\eta: \mathcal{A}(G) \rightarrow \mathcal{A}({}^{\tau,x}G) \tag{22}$$

compatible with (21), constructed in [36, Lemma 8.2].

**Proposition 5.3.** *For every finite prime  $v$  of  $F$ , the quaternion algebra  ${}^{\tau,x}B$  splits over  $v$  if and only if  $B$  splits over  $v$ . The set of infinite primes of  $F$  where  ${}^{\tau,x}B$  splits is equal to  $\tau\mathcal{P}$ .*

*Proof.* This is, in a slightly different language, the main result of [46]. It is stated in the language we use (but without proof and with a sign error) in [34, Remark 4.1(b)]. Note that the statement about the finite primes follows directly from (20); for the infinite primes some tedious but elementary calculations are needed. We refer to [46] for these, but it should be possible to replace them by the explicit cohomological calculations of [37].  $\square$

*The twisted Shimura datum.* Let  ${}^{\tau,x}G$  be the semisimple group constructed before, and let  ${}^{\tau,x}G^{\text{ad}}$  be its adjoint group. Note that the torus  ${}^{\tau,x}T \subset {}^{\tau,x}G^{\text{ad}}$  is canonically isomorphic to  $T$  (even though the cohomology class  $t_\tau \in H^1(\mathbb{Q}, T)$  may be non-trivial) because the action of  $T$  on itself by inner automorphisms is trivial. Hence we may apply  $\tau$  to the cocharacter  $\mu_x: \mathbb{G}_{m,\mathbb{C}} \rightarrow T_{\mathbb{C}}$  and interpret the result as a cocharacter  $\tau\mu_x: \mathbb{G}_{m,\mathbb{C}} \rightarrow {}^{\tau,x}T_{\mathbb{C}}$ . This gives rise to a homomorphism  ${}^{\tau}h: \mathbb{S} \rightarrow {}^{\tau,x}G_{\mathbb{R}}^{\text{ad}}$ . We let  ${}^{\tau,x}X$  be the  ${}^{\tau,x}G^{\text{ad}}(\mathbb{R})^+$ -conjugacy class of homomorphisms  $\mathbb{S} \rightarrow {}^{\tau,x}G_{\mathbb{R}}^{\text{ad}}$  containing  ${}^{\tau}h$ . We denote  ${}^{\tau}h$ , when viewed as a point in  ${}^{\tau,x}X$ , by  ${}^{\tau}x$ . An easy calculation shows that then

$${}^{\tau,x}(G, X) = ({}^{\tau,x}G, {}^{\tau,x}X)$$

is again a connected Shimura datum.

Its ‘classical’ description is this: for every  $\varrho \in \mathcal{P}$  we are given an isomorphism  ${}^{\tau,x}B \otimes_{F,\tau\varrho} \mathbb{R} \simeq M_2(\mathbb{R})$ , and hence an operation of  ${}^{\tau,x}G^{\text{ad}}(\mathbb{R})$  on  $\mathbb{H}^{\tau\mathcal{P}}$  by coordinate-wise Möbius transformations; corresponding to these isomorphisms we also obtain an isomorphism  ${}^{\tau,x}X \simeq \mathbb{H}^{\tau\mathcal{P}}$ .

We are now ready to describe exactly how  $\tau$  acts on the quaternionic Shimura varieties we consider.

**Theorem 5.4.** *There exists a unique isomorphism of  $\mathbb{C}$ -schemes*

$$\varphi_{\tau,x}^0: \tau\text{Sh}^0(G, X) \rightarrow \text{Sh}^0({}^{\tau,x}G, {}^{\tau,x}X)$$

which takes  $\tau[x]$  to  $[{}^{\tau}x]$  and is equivariant for

$$\eta: \mathcal{A}(G) \rightarrow \mathcal{A}({}^{\tau,x}G)$$

as in (22).

This is a special case of a result of Milne–Shih [32], conjectured by Langlands [28]. For this case this result had been proved earlier in more classical language in [15] for Shimura curves and in [46] for quaternionic Shimura varieties of arbitrary dimension. The present author prefers the more abstract setup utilised here because it also leads to a useful description of the Galois actions on automorphic bundles, needed later on.

Let  $\Gamma \subset G(\mathbb{Q}) = B^1$  be a congruence subgroup, let  $\hat{\Gamma}$  be its congruence completion, i.e. its topological closure in  $G(\mathbb{A}^f)$ , and let  ${}^{\tau,x}\Gamma = {}^{\tau,x}G(\mathbb{Q}) \cap \eta(\hat{\Gamma})$ . Then  ${}^{\tau,x}\Gamma$  is a congruence subgroup of  ${}^{\tau,x}G(\mathbb{Q}) \simeq (B')^1$ , and we obtain an isomorphism of complex algebraic varieties

$$(\varphi_{\tau,x}^0)_\Gamma: \tau\text{Sh}_\Gamma^0(G, X) \rightarrow \text{Sh}_{{}^{\tau,x}\Gamma}^0({}^{\tau,x}G, {}^{\tau,x}X).$$

*Example 5.5.* (i) Let  $\mathcal{O} \subset B$  be an order, and let  $\hat{\mathcal{O}} = \mathcal{O} \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}$  be its closure in  $B_{\mathbb{A}^f} = B \otimes_F \mathbb{A}_F^f$ . Then  $\hat{\mathcal{O}}$  is a compact open subring of  $B_{\mathbb{A}^f}$ , hence  $\eta(\hat{\mathcal{O}})$  is a compact open subring of  ${}^{\tau,x}B_{\mathbb{A}^f}$ . We set

$${}^{\tau,x}\mathcal{O} \stackrel{\text{def}}{=} \eta(\hat{\mathcal{O}}) \cap {}^{\tau,x}B.$$

This is an order in  ${}^{\tau,x}B$ ; it is maximal if and only if  $\mathcal{O}$  is maximal.

Then  $\Gamma = \mathcal{O}^1$  is a congruence subgroup of  $G(\mathbb{Q})$  and  ${}^{\tau,x}\Gamma = ({}^{\tau,x}\mathcal{O})^1$  is a congruence subgroup of  ${}^{\tau,x}G(\mathbb{Q})$ . If  $S$  is the complex algebraic variety with  $S(\mathbb{C}) = \mathcal{O}^1 \backslash \mathbb{H}^{\mathcal{P}}$  and  ${}^{\tau,x}S$  is the complex algebraic variety with  ${}^{\tau,x}S(\mathbb{C}) =$

$({}^{\tau,x}\mathcal{O})^1 \backslash \mathbb{H}^{\tau\mathcal{P}}$ , then the above-mentioned theorem states the existence of a canonical isomorphism  $\tau S \cong {}^{\tau,x}S$ .

- (ii) Continue with the notation from the previous example, and let  $\mathfrak{a}$  be an integral ideal of  $F$ . Using the fact that  $\eta$  is an isomorphism of  $\mathbb{A}_F^f$ -algebras we find the formula

$${}^{\tau,x}(\Gamma(\mathfrak{a})) = ({}^{\tau,x}\Gamma)(\mathfrak{a})$$

for principal congruence subgroups. Hence if we let  $S(\mathfrak{a})$  and  ${}^{\tau,x}S(\mathfrak{a})$  be the varieties with  $S(\mathfrak{a})(\mathbb{C}) = \mathcal{O}^1(\mathfrak{a}) \backslash X$  and  ${}^{\tau,x}S(\mathfrak{a})(\mathbb{C}) = ({}^{\tau,x}\mathcal{O})^1(\mathfrak{a}) \backslash {}^{\tau,x}X$ , then  $\tau S(\mathfrak{a}) \cong {}^{\tau,x}S(\mathfrak{a})$ .

*Remark 5.6.* (i) The point  $x$  is always a special point, but only a CM point if  $B$  is totally indefinite. This is not a defect of our construction but necessary: CM points exist only if the weight homomorphism  $\mathbb{G}_{m,\mathbb{R}} \rightarrow \mathbb{S} \rightarrow G_{\mathbb{R}}^{\text{ad}}$  is defined over  $\mathbb{Q}$  (compare the discussion in [35, Section 12]). In our case the field of definition of the weight homomorphism (which is independent of  $x$ ) is the fixed field in  $\overline{\mathbb{Q}}$  of

$$\{\sigma \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \mid \sigma \circ \mathcal{P} = \mathcal{P}\}.$$

Clearly this is equal to  $\mathbb{Q}$  if and only if  $\mathcal{P} = \mathcal{S}_{\infty}$ , the set of all archimedean places of  $F$ .

If this is the case, the map  $\varrho_x: \mathfrak{S} \rightarrow T$  can be lifted to  $\mathbb{T}^E$  and the above construction simplifies considerably. The cohomology class  $\varrho_{x,*}(s_{\tau})$  is then the image of a class in  $H^1(\mathbb{Q}, \mathbb{T}^E) \cong H^1(E, \mathbb{G}_m) = 0$  and therefore the connected Shimura datum  ${}^{\tau,x}(G, X)$  is isomorphic to  $(G, X)$ .

- (ii) The choice of  $x$  is inessential: changing it amounts to applying certain canonical isomorphisms everywhere, cf. [33, Proposition 1.3 and Lemma 5.1].

*Galois conjugates of automorphic bundles.* Let  ${}^{\tau,x}\check{X}$  be the compact dual symmetric space of  ${}^{\tau,x}X$ . This comes with a natural action of  ${}^{\tau,x}G_{\mathbb{C}}$ . Since  ${}^{\tau,x}X$  is derived from a quaternion algebra  ${}^{\tau,x}B$  over  $F$ , we can construct a  ${}^{\tau,x}G_{\mathbb{C}}$ -equivariant line bundle  ${}^{\tau,x}\mathfrak{L}_{\varrho}$  on  ${}^{\tau,x}\check{X}$ , for every  $\varrho: F \rightarrow \mathbb{R}$  at which  ${}^{\tau,x}B$  is unramified; that is, for every  $\varrho \in \tau\mathcal{P}$ .

Now the trivialisation  $z_{\infty}(\tau) \in {}^{\tau}\mathfrak{S}(\mathbb{C})$  defines an isomorphism of complex algebraic groups

$$G_{\mathbb{C}} \rightarrow {}^{\tau,x}G_{\mathbb{C}}, \quad g \mapsto [z_{\infty}(\tau), g]. \quad (23)$$

**Proposition 5.7.** *There is a commutative diagram*

$$\begin{array}{ccc} {}^{\tau}\mathfrak{L}_{\varrho} & \xrightarrow{\cong} & {}^{\tau,x}\mathfrak{L}_{\tau \circ \varrho} \\ \downarrow & & \downarrow \\ {}^{\tau}\check{X} & \xrightarrow[\cong]{} & {}^{\tau,x}\check{X} \end{array}$$

where the vertical maps are the natural structure maps of line bundles, the horizontal isomorphisms are equivariant for (23) and the lower horizontal map sends  $\tau[x]$  to  $[\tau x]$ .

*Proof.* This is a simple combination of [33, Proposition 2.7 and Theorem 3.10], but in our special case it can also be checked by a straightforward calculation.  $\square$

The  ${}^{\tau,x}G_{\mathbb{C}}$ -line bundle  ${}^{\tau,x}\mathfrak{L}_{\varrho}$  on  ${}^{\tau,x}\check{X}$  gives rise to an automorphic line bundle  ${}^{\tau,x}\mathcal{L}_{\varrho}$  on  $\mathrm{Sh}^0({}^{\tau,x}G, {}^{\tau,x}X)$ , which comes with a continuous  ${}^{\tau,x}G(\mathbb{A}^f)$ -action, just as for  $\mathfrak{L}_{\varrho}$  on  $\check{X}$ .

**Theorem 5.8.** *The exists an isomorphism of line bundles  $\tau\mathcal{L}_{\varrho} \rightarrow {}^{\tau,x}\mathcal{L}_{\tau\varrho}$  covering the isomorphism  $\varphi_{\tau,x}^0: \tau\mathrm{Sh}^0(G, X) \rightarrow \mathrm{Sh}^0({}^{\tau,x}G, {}^{\tau,x}X)$  and equivariant for  $\eta: G(\mathbb{A}^f) \rightarrow {}^{\tau,x}G(\mathbb{A}^f)$ .*

*Proof.* This is a special case of Milne's result [33, Theorem 5.2].  $\square$

**5.4. Galois conjugates of modular embeddings.** In this subsection we prove Theorem B, first in the simple version stated in the introduction (Theorem 5.9 below) and then in a more elaborate version (Theorem 5.12 below) taking into account the adelic structures.

**Theorem 5.9.** *Let  $C$  be a smooth complex curve, let  $S = \mathrm{Sh}_{\Gamma}^0(G, X)$  be a quaternionic Shimura variety, with  $\Gamma$  torsion-free. Let  $f: C \rightarrow S$  be covered by a modular embedding with respect to  $\varrho: F \rightarrow \mathbb{R}$ . Let  $\tau \in \mathrm{Aut} \mathbb{C}$ , and identify  $\tau S$  with a quaternionic Shimura variety via the isomorphism  $(\varphi_{\tau,x}^0)_{\Gamma}$ .*

*Then  $\tau f: \tau C \rightarrow \tau S$  is covered by a modular embedding with respect to  $\tau \circ \varrho$ .*

*Proof.* For projective  $C$  this follows easily by combining Theorems 4.4 and 5.8.

For affine  $C$  we use Theorem 4.8 instead. Note that all constructions involved are purely algebraic and stable under field automorphisms: relative de Rham cohomology and the Gauss–Manin connection are, and the Deligne extension of relative de Rham cohomology can be characterised as the unique extension to an algebraic vector bundle with algebraic logarithmic connection having nilpotent residue, cf. Remark 2.4.  $\square$

*Field automorphisms and uniformisation for curves.* Let  $\Delta \subset \mathrm{PGL}_2(\mathbb{R})^+$  be a cocompact lattice, and let  $\tau \in \mathrm{Aut} \mathbb{C}$ . Then there exists a lattice  ${}^{\tau}\Delta \subset \mathrm{PGL}_2(\mathbb{R})^+$ , unique up to  $\mathrm{PGL}_2(\mathbb{R})^+$ -conjugacy, such that applying  $\tau$  to the complex algebraic curve underlying  $\Delta \backslash \mathbb{H}$  yields a curve isomorphic to  ${}^{\tau}\Delta \backslash \mathbb{H}$ , where the isomorphism is in addition required to preserve the orders of the elliptic points. This is of course a simple construction, but for our purposes there are two disadvantages to this elementary formulation: the group  ${}^{\tau}\Delta$  is only well-defined up to conjugacy, and it does not harmonise with the more elaborate constructions needed to describe Galois conjugates of Shimura varieties.

Therefore we fix a point  $y \in \mathbb{H}$ . We shall construct an hermitian symmetric domain  ${}^{\tau,y}\mathbb{H}$ , a real algebraic group  ${}^{\tau,y}\mathrm{PGL}_2$  with an action of  ${}^{\tau,y}\mathrm{PGL}_2(\mathbb{R})^+$  on  ${}^{\tau,y}\mathbb{H}$  by biholomorphisms, and a lattice  ${}^{\tau,y}\Delta \subset {}^{\tau,y}\mathrm{PGL}_2(\mathbb{R})^+$ .

There is a well-defined smooth curve  $C$  over  $\mathbb{C}$  with  $C(\mathbb{C}) \cong \Delta \backslash \mathbb{H}$ , and there is a map  $e: C(\mathbb{C}) \rightarrow \mathbb{N}$  sending the image of a point  $x \in \mathbb{H}$  to the order of the stabiliser  $\Delta_x$ . Let  $c \in C(\mathbb{C})$  be the image of  $y \in \mathbb{H}$ . Then  $(\tau C, \tau c)$  is again a smooth complex curve equipped with a geometric point, and we define  $({}^{\tau,y}\mathbb{H}, {}^{\tau}y) \rightarrow ((\tau C)^{\mathrm{an}}, \tau c)$  to be a ramified covering space which has ramification order  $e(\tau^{-1}(p))$  at  $p \in \tau C(\mathbb{C})$  and which is universal with these properties. Then  ${}^{\tau,y}\mathbb{H}$  inherits a natural complex structure turning it into a bounded symmetric domain, and the Bergman metric is a canonical hermitian metric on  ${}^{\tau,y}\mathbb{H}$  turning it into a hermitian symmetric domain.

It is isomorphic, *but not canonically*, to the upper half plane. The group  $B$  of biholomorphisms of  ${}^{\tau,y}\mathbb{H}$  is a real Lie group, and we let  $\mathfrak{b}$  be its Lie algebra. Then there exists a unique connected real algebraic subgroup  ${}^{\tau,y}\mathrm{PGL}_2 \subset \mathrm{GL}(\mathfrak{b})$  such that  ${}^{\tau,y}\mathrm{PGL}_2(\mathbb{R})^+ = \mathrm{Ad} B \cong B$  in  $\mathrm{GL}(\mathfrak{b})$ . By construction, the deck transformation group  ${}^{\tau,y}\Delta$  of  ${}^{\tau,y}\mathbb{H} \rightarrow (\tau C)^{\mathrm{an}}$  is a lattice in  ${}^{\tau,y}\mathrm{PGL}_2(\mathbb{R})^+ = B$ , and there is a natural isomorphism  $\tau(\Delta \backslash \mathbb{H}) \cong {}^{\tau,y}\Delta \backslash {}^{\tau,y}\mathbb{H}$  preserving the orders of elliptic points.

**Proposition 5.10.** *The groups  $\Delta$  and  ${}^{\tau,y}\Delta$  are isomorphic as abstract groups. The isomorphism can be chosen in such a way as to send elliptic, parabolic and hyperbolic elements to elements of the same type.*

*Proof.* The isomorphism type of  $\Delta$  is uniquely determined by the orders of elliptic points, the number of cusps and the genus of the quotient. These are the same as for  ${}^{\tau,y}\Delta$ . Elliptic, parabolic and hyperbolic elements can be characterised in purely topological terms.  $\square$

There seems to be no canonical way to construct such an isomorphism  $\Delta \rightarrow {}^{\tau,y}\Delta$  for all lattices  $\Delta$ . Note, however, that the profinite completions of  $\Delta$  and  ${}^{\tau,y}\Delta$  are canonically isomorphic since they can be interpreted as étale fundamental groups. In Theorem 5.12 below we will see that if  $\Delta$  admits a modular embedding then so does  ${}^{\tau,y}\Delta$ , and there is a canonical isomorphism between their congruence completions.

*Modular embeddings, once more.* Suppose there is given a modular embedding consisting of the following data: a lattice  $\Delta \subset \mathrm{PGL}_2(\mathbb{R})^+$ , a quaternionic connected Shimura datum  $(G, X)$ , an arithmetic subgroup  $\Gamma \subset G^{\mathrm{ad}}(\mathbb{Q})^+$  containing a group in  $\Sigma(G)$ , a homomorphism  $\varphi: \Delta \rightarrow \Gamma$  and a  $\varphi$ -equivariant holomorphic map  $\tilde{f}: \mathbb{H} \rightarrow X$ .

As before, we obtain a regular map  $f: C \rightarrow S$  between normal varieties with  $C^{\mathrm{an}} = \Delta \backslash \mathbb{H}$  and  $S = \mathrm{Sh}_\Gamma^0(G, X)$ , and congruence completions  $\hat{\varphi}: \hat{\Delta} \rightarrow \mathcal{A}(G)$  and  $\hat{f}: \hat{C} \rightarrow \mathrm{Sh}^0(G, X)$ .

To formulate the next theorem, we need to assume one condition:

*Condition 5.11.* There exists a point  $y \in \mathbb{H}$  such that  $x = \tilde{f}(y) \in X$  is a special point for  $\Gamma$ .

This is satisfied, for example, if  $y$  is an elliptic fixed point for some group commensurable to  $\Delta$ . In particular, for  $\Delta$  commensurable to a triangle group we can always find such a  $y$ . On the other hand, the André–Oort conjecture for  $S$  predicts that as soon as  $\Delta$  is nonarithmetic there are at most finitely many possible  $\Delta$ -orbits for  $y$ .

**Theorem 5.12.** *Let  $y \in \mathbb{H}$  be a point as in Condition 5.11.*

- (i) *There exists a modular embedding for  ${}^{\tau,y}\Delta$  relative to the field embedding  $\tau \circ \varrho$ , consisting of a group homomorphism  ${}^{\tau,y}\varphi: {}^{\tau,y}\Delta \rightarrow {}^{\tau,x}\Gamma \subset {}^{\tau,x}G^{\mathrm{ad}}(\mathbb{Q})^+$  and a  ${}^{\tau,y}\varphi$ -equivariant holomorphic map  ${}^{\tau,y}\tilde{f}: {}^{\tau,y}\mathbb{H} \rightarrow {}^{\tau,x}X$ .*

(ii) *There exists a commutative diagram of  $\mathbb{C}$ -schemes*

$$\begin{array}{ccc} \tau\hat{C} & \xrightarrow{\tau\hat{f}} & \tau\mathrm{Sh}^0(G, X) \\ \downarrow \cong & & \downarrow \cong \\ \widehat{\tau,yC} & \xrightarrow[\tau,yf]{} & \mathrm{Sh}^0(\tau,xG, \tau,xX) \end{array}$$

*equivariant for the commutative diagram of topological groups*

$$\begin{array}{ccc} \hat{\Delta} & \xrightarrow{\hat{\varphi}} & \mathcal{A}(G) \\ \downarrow \cong & & \downarrow \cong \eta \\ \widehat{\tau,y\Delta} & \xrightarrow[\tau,y\varphi]{} & \mathcal{A}(\tau,xG). \end{array}$$

*Proof.* The morphism  $\tau f: \tau C \rightarrow \tau S$  admits a unique holomorphic lift to the simply-connected covering spaces

$$\tau,y\tilde{f}: \tau,y\mathbb{H} \rightarrow \tau,yX$$

such that  $\tau,y\tilde{f}(\tau y) = \tau x$ . It is equivariant for a homomorphism of deck transformation groups

$$\tau,y\varphi: \tau,y\Delta \rightarrow \tau,x\Gamma.$$

By Theorem 5.9 applied to suitable finite-index subgroups,  $\tau,y\tilde{f}$  and  $\tau,y\varphi$  constitute a modular embedding.

It remains to construct the isomorphisms  $\tau\hat{C} \rightarrow \widehat{\tau,yC}$  and  $\hat{\Delta} \rightarrow \widehat{\tau,y\Delta}$ . To do this, note that a covering  $C' \rightarrow C$  coming from a finite-index subgroup  $\Delta' \subset \Delta$  is a congruence covering if and only if it arises as the pullback along  $f: C \rightarrow S$  of a congruence covering of  $S$ . Since the latter are stable under  $\tau$ , congruence coverings of  $C$  correspond bijectively to congruence covers of  $\tau C = \tau,yC$ , and we obtain the desired isomorphisms.  $\square$

*Congruence subgroups.* Assume now in addition that (12) holds. Recall that both triangle groups and fundamental groups of Teichmüller curves admit modular embeddings satisfying (12). Then  $\varphi$  can be lifted to a homomorphism  $\tilde{\varphi}: \tilde{\Delta} \rightarrow G(\mathbb{Q})$  where  $\tilde{\Delta} \subset \mathrm{SL}_2(\mathbb{R})$  is the preimage of  $\Delta$  and where

$$\mathrm{tr} \tilde{\delta} = \varrho(\mathrm{tr} \tilde{\varphi}(\tilde{\delta})) \quad \text{for all } \tilde{\delta} \in \tilde{\Delta}.$$

By Proposition 4.2.(ii) the trace field  $K = \mathbb{Q}(\mathrm{tr} \Delta)$  is then contained in  $\varrho(F)$ . By restricting to a subalgebra of  $B$  we will assume in addition that  $K = \varrho(F)$ ; this is an inessential restriction but it leads to a simpler notation. We obtain a quaternion algebra  $A = K\langle\tilde{\Delta}\rangle \subset M_2(\mathbb{R})$  over  $K$  and an order  $\mathcal{O} = \mathfrak{o}_K\langle\tilde{\Delta}\rangle$  of  $A$ . The map  $\tilde{\varphi}$  extends to an isomorphism  $A \rightarrow B$  of  $\mathbb{Q}$ -algebras which is semilinear for  $(\varrho|_K)^{-1}: K \rightarrow F$ , and it sends  $\mathcal{O}$  to an order  $\mathcal{O}_B$  of  $B$ .

We may then set  $\tilde{\Gamma} = \mathcal{O}_B^1$  and let  $\Gamma$  be its image in  $G^{\mathrm{ad}}(\mathbb{Q})^+$ . Then  $\varphi$  may be considered as a homomorphism  $\varphi: \Delta \rightarrow \Gamma$ . For every ideal  $\mathfrak{n}$  of  $F$  we set

$$\tilde{\Gamma}(\mathfrak{n}) = \{\tilde{\gamma} \in \tilde{\Gamma} \mid \tilde{\gamma} - 1 \in \mathfrak{n}\mathcal{O}_B\},$$

and for an ideal  $\mathfrak{a}$  of  $K$  we set

$$\tilde{\Delta}(\mathfrak{a}) = \{\tilde{\delta} \in \tilde{\Gamma} \mid \tilde{\delta} - 1 \in \mathfrak{a}\mathcal{O}\} = \tilde{\varphi}^{-1}(\tilde{\Gamma}(\varrho^{-1}\mathfrak{a})).$$

We let  $\Gamma(\mathfrak{n})$  and  $\Delta(\mathfrak{a})$  be the images of these groups  $G^{\text{ad}}(\mathbb{Q})^+$  and  $\text{PGL}_2(\mathbb{R})^+$ , respectively.

**Corollary 5.13.** *Let  $\Delta \subset \text{PGL}_2(\mathbb{R})^+$  be a lattice admitting a modular embedding with (12) and with trace field  $K$ . Let  $y \in \mathbb{H}$  and  $\tau \in \text{Aut } \mathbb{C}$ , and let  $\mathfrak{a}$  be an ideal in  $K$ . Then the trace field of  ${}^{\tau,y}\Delta$  is  $\tau(K)$ , and  ${}^{\tau,y}(\Delta(\mathfrak{a})) = ({}^{\tau,y}\Delta)(\tau\mathfrak{a})$ .*

*Proof.* An easy argument shows that  $\hat{\mathcal{O}} = \hat{\mathfrak{o}}_F \langle \hat{\varphi}(\hat{\Delta}) \rangle$  and therefore  ${}^{\tau,x}\hat{\mathcal{O}} = \hat{\mathfrak{o}}_F \langle \widehat{{}^{\tau,y}\varphi({}^{\tau,y}\Delta)} \rangle$ . This implies  $\mathfrak{o}_F \langle {}^{\tau,y}\varphi({}^{\tau,y}\Delta) \rangle$  is dense in  ${}^{\tau,x}\hat{\mathcal{O}}$ ; since an order in a quaternion algebra is uniquely determined by its adelic closure we obtain

$${}^{\tau,x}\mathcal{O} = \mathfrak{o}_K \langle {}^{\tau,y}\varphi({}^{\tau,y}\Delta) \rangle.$$

Now  $\widehat{\Delta(\mathfrak{a})} = \widehat{\hat{\varphi}^{-1}(\Gamma(\mathfrak{a}))}$  and

$${}^{\tau,y}\widehat{\Delta(\mathfrak{a})} = \widehat{{}^{\tau,y}\varphi^{-1}(\Gamma(\mathfrak{a}))} \stackrel{(*)}{=} \widehat{{}^{\tau,y}\varphi^{-1}(({}^{\tau,x}\Gamma)(\mathfrak{a}))} = \widehat{({}^{\tau,y}\Delta)(\tau\mathfrak{a})},$$

where  $(*)$  is justified by Example 5.5. Since  $\Delta$  and  ${}^{\tau,y}\Delta$  are dense in their respective congruence completions, this implies that  ${}^{\tau,y}(\Delta(\mathfrak{a})) = ({}^{\tau,y}\Delta)(\tau\mathfrak{a})$ .  $\square$

**5.5. Application to triangle groups and dessins d'enfants.** Let  $\Delta_{p,q,r}$  be a Fuchsian triangle group, let  $\mathfrak{a}$  be an ideal in its trace field and let  $X_{p,q,r}(\mathfrak{a})$  be the smooth projective complex curve with associated Riemann surface  $\Delta_{p,q,r}(\mathfrak{a}) \backslash \mathbb{H}$ . There is a unique isomorphism  $X_{p,q,r}(1) \cong \mathbb{P}^1$  sending the elliptic points of order  $p, q, r$  to the points  $0, 1, \infty$ , respectively. Hence the quotient map  $X_{p,q,r}(\mathfrak{a}) \rightarrow X_{p,q,r}(1) \cong \mathbb{P}^1$  is a normal *Belyi map*: it is unramified outside  $0, 1, \infty$ .

Belyi maps are famously in one-to-one correspondence with the combinatorial-topological objects known as *dessins d'enfants*; the Galois action on dessins d'enfants is much studied for its importance in anabelian geometry. Let  $\mathcal{D}_{p,q,r}(\mathfrak{a})$  be the dessin associated with the aforementioned Belyi map.

**Theorem 5.14.** *The action of  $\text{Aut } \mathbb{C}$  on the curves  $X_{p,q,r}(\mathfrak{a})$  and the dessins  $\mathcal{D}_{p,q,r}(\mathfrak{a})$  factors through  $\text{Gal}(K_{p,q,r}/\mathbb{Q})$ ; it is given by*

$$\tau X_{p,q,r}(\mathfrak{a}) \cong X_{p,q,r}(\tau\mathfrak{a}) \quad \text{and} \quad \tau \mathcal{D}_{p,q,r}(\mathfrak{a}) \cong \mathcal{D}_{p,q,r}(\tau\mathfrak{a}).$$

*Proof.* The triangle group  $\Delta_{p,q,r}$  is a lattice in  $\text{PGL}_2(\mathbb{R})^+$  with presentation

$$\Delta_{p,q,r} = \langle x, y, z \mid x^p = y^q = z^r = xyz = 1 \rangle.$$

Moreover, it is the only lattice in  $\text{PGL}_2(\mathbb{R})^+$  with that presentation, up to conjugacy. Therefore if  $\tau \in \text{Aut } \mathbb{C}$  and  $y \in \mathbb{H}$ , and if we choose a biholomorphism  ${}^{\tau,y}\mathbb{H} \cong \mathbb{H}$  and hence an isomorphism  ${}^{\tau,y}\text{PGL}_{2,\mathbb{R}} \cong \text{PGL}_{2,\mathbb{R}}$ , then by Proposition 5.10  ${}^{\tau,y}\Delta_{p,q,r}$  must be conjugate in  $\text{PGL}_2(\mathbb{R})^+$  to  $\Delta_{p,q,r}$ , and we may assume the identifications are chosen in such a way that

$${}^{\tau,y}\Delta_{p,q,r} = \Delta_{p,q,r}.$$

By Corollary 5.13 we obtain

$${}^{\tau,y}(\Delta_{p,q,r}(\mathfrak{a})) = \Delta_{p,q,r}(\tau\mathfrak{a}). \quad \square$$

Special cases of this theorem were known before. In [18, Satz 2.5.1] it was proved for signatures of the form  $(2, q, r)$  satisfying some additional regularity property and for prime ideals  $\mathfrak{a}$ . Furthermore it was proved for all cocompact arithmetic triangle groups except  $\Delta_{3,4,6}$  in [20, Theorem 6]. Note that this is only contained in the preprint version [20], not in the published version [19]. These two works used a different method, analysing the operation of  $\Delta/\Delta(\mathfrak{a})$  on the space of holomorphic one-forms on  $X(\mathfrak{a})$ .

The groups  $\Delta_{p,q,r}(\mathfrak{a})$  and the curves  $X_{p,q,r}(\mathfrak{a})$  are studied in detail in [4]. In [27] the arithmetic aspects of modular embeddings for triangle groups and their congruence subgroups will be studied further.

## REFERENCES

- [1] Armand Borel, *Introduction aux groupes arithmétiques*, Publications de l'Institut de Mathématique de l'Université de Strasbourg, XV. Actualités Scientifiques et Industrielles, No. 1341, Hermann, Paris, 1969.
- [2] ———, *Some metric properties of arithmetic quotients of symmetric spaces and an extension theorem*, J. Differential Geometry **6** (1972), pp. 543–560, Collection of articles dedicated to S. S. Chern and D. C. Spencer on their sixtieth birthdays.
- [3] Armand Borel and Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. of Math. (2) **75** (1962), pp. 485–535.
- [4] Pete L. Clark and John Voight, *Algebraic curves uniformized by congruence subgroups of triangle groups*, 2015, arXiv:1506.01371.
- [5] Paula Cohen and Jürgen Wolfart, *Modular embeddings for some nonarithmetic Fuchsian groups*, Acta Arith. **56** (1990), no. 2, pp. 93–110.
- [6] Marc Culler, *Lifting representations to covering groups*, Adv. in Math. **59** (1986), no. 1, pp. 64–70.
- [7] Pierre Deligne, *Équations différentielles à points singuliers réguliers*, Lecture Notes in Mathematics, Vol. 163, Springer-Verlag, Berlin-New York, 1970.
- [8] ———, *Théorie de Hodge. II*, Inst. Hautes Études Sci. Publ. Math. (1971), no. 40, pp. 5–57.
- [9] ———, *Travaux de Shimura*, Séminaire Bourbaki, 23ème année (1970/71), Exp. No. 389, Springer, Berlin, 1971, pp. 123–165. Lecture Notes in Math., Vol. 244.
- [10] ———, *Cohomologie étale*, Lecture Notes in Mathematics, Vol. 569, Springer-Verlag, Berlin-New York, 1977, Séminaire de Géométrie Algébrique du Bois-Marie SGA 4 $\frac{1}{2}$ , avec la collaboration de J. F. Boutot, A. Grothendieck, L. Illusie et J. L. Verdier.
- [11] ———, *Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques*, Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, pp. 247–289.
- [12] ———, *Motifs et groupe de Taniyama*, Hodge cycles, Motives, and Shimura varieties, Lecture notes in Mathematics, vol. 900, Springer, Berlin, Heidelberg, 1982, pp. 261–279.
- [13] Pierre Deligne and James S. Milne, *Tannakian categories*, Hodge cycles, Motives, and Shimura varieties, Lecture notes in Mathematics, vol. 900, Springer, Berlin, Heidelberg, 1982, pp. 101–228.
- [14] Pierre Deligne (Notes by J. Milne), *Hodge cycles on abelian varieties*, Hodge cycles, Motives, and Shimura varieties, Lecture notes in Mathematics, vol. 900, Springer, Berlin, Heidelberg, 1982, pp. 101–228.
- [15] Koji Doi and Hidehisa Naganuma, *On the algebraic curves uniformized by arithmetical automorphic functions*, Ann. of Math. (2) **86** (1967), pp. 449–460.
- [16] Michel Emsalem, *Twisting by a torsor*, 2015, arXiv:1508.02903.
- [17] Gerd Faltings, *Real projective structures on Riemann surfaces*, Compositio Math. **48** (1983), no. 2, pp. 223–269.

- [18] Frank Feierabend, *Galois-Operationen auf verallgemeinerten Macbeath-Hurwitz-Kurven*, 2008, PhD dissertation, Frankfurt.
- [19] Ernesto Gironde, David Torres-Teigell, and Jürgen Wolfart, *Fields of definition of uniform dessins on quasisplatonic surfaces*, Riemann and Klein surfaces, automorphisms, symmetries and moduli spaces, Contemp. Math., vol. 629, Amer. Math. Soc., Providence, RI, 2014, pp. 155–170.
- [20] ———, *Fields of definition of uniform dessins on quasisplatonic surfaces (extended version)*, 2014, available at <http://www.uni-frankfurt.de/50936228/Publikationen>.
- [21] Phillip A. Griffiths, *Periods of integrals on algebraic manifolds. III. Some global differential-geometric properties of the period mapping*, Inst. Hautes Études Sci. Publ. Math. (1970), no. 38, pp. 125–180.
- [22] R. C. Gunning, *Special coordinate coverings of Riemann surfaces*, Math. Ann. **170** (1967), pp. 67–86.
- [23] F. Hirzebruch and D. Zagier, *Intersection numbers of curves on Hilbert modular surfaces and modular forms of Nebentypus*, Invent. Math. **36** (1976), pp. 57–113.
- [24] Irwin Kra, *On lifting Kleinian groups to  $SL(2, \mathbb{C})$* , Differential geometry and complex analysis, Springer, Berlin, 1985, pp. 181–193.
- [25] Robert A. Kucharczyk, *Real Multiplication on Jacobian Varieties*, 2012, arXiv:1201.1570.
- [26] ———, *Modular embeddings and rigidity for Fuchsian groups*, Acta Arith. **169** (2015), no. 1, pp. 77–100.
- [27] Robert A. Kucharczyk and John Voight, *Canonical models for triangle modular curves*, in preparation.
- [28] Robert P. Langlands, *Automorphic representations, Shimura varieties, and motives. Ein Märchen*, Automorphic forms, representations and  $L$ -functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, pp. 205–246.
- [29] Joseph Le Potier, *Fibrés de Higgs et systèmes locaux*, Astérisque (1991), no. 201-203, Exp. No. 737, pp. 221–268 (1992), Séminaire Bourbaki, Vol. 1990/91.
- [30] Colin Maclachlan and Alan W. Reid, *The arithmetic of hyperbolic 3-manifolds*, Graduate Texts in Mathematics, vol. 219, Springer-Verlag, New York, 2003.
- [31] Curtis T. McMullen, *Billiards and Teichmüller curves on Hilbert modular surfaces*, J. Amer. Math. Soc. **16** (2003), no. 4, pp. 857–885.
- [32] James S. Milne, *The action of an automorphism of  $\mathbb{C}$  on a Shimura variety and its special points*, Arithmetic and geometry, Vol. I, Progr. Math., vol. 35, Birkhäuser Boston, Boston, MA, 1983, pp. 239–265.
- [33] ———, *Automorphic vector bundles on connected Shimura varieties*, Invent. Math. **92** (1988), no. 1, pp. 91–128.
- [34] ———, *Canonical models of (mixed) Shimura varieties and automorphic vector bundles*, Automorphic forms, Shimura varieties, and  $L$ -functions, Vol. I (Ann Arbor, MI, 1988), Perspect. Math., vol. 10, Academic Press, Boston, MA, 1990, pp. 283–414.
- [35] ———, *Introduction to Shimura varieties*, Harmonic analysis, the trace formula, and Shimura varieties, Clay Math. Proc., vol. 4, Amer. Math. Soc., Providence, RI, 2005, pp. 265–378.
- [36] James S. Milne and Kuang-ye Shih, *Conjugates of Shimura varieties*, Hodge cycles, Motives, and Shimura varieties, Lecture notes in Mathematics, vol. 900, Springer, Berlin, Heidelberg, 1982, pp. 280–356.
- [37] ———, *Langlands’s construction of the Taniyama group*, Hodge cycles, Motives, and Shimura varieties, Lecture notes in Mathematics, vol. 900, Springer, Berlin, Heidelberg, 1982, pp. 229–260.
- [38] Martin Möller, *Variations of Hodge structures of a Teichmüller curve*, J. Amer. Math. Soc. **19** (2006), no. 2, pp. 327–344.
- [39] Martin Möller and Eckart Viehweg, *Kobayashi geodesics in  $\mathcal{A}_g$* , J. Differential Geom. **86** (2010), no. 2, pp. 355–579.
- [40] Martin Möller and Don Zagier, *Modular embeddings of Teichmüller curves*, 2015, arXiv:1503.05690.

- [41] Wilfried Schmid, *Variation of Hodge structure: the singularities of the period mapping*, Invent. Math. **22** (1973), pp. 211–319.
- [42] Paul Schmutz Schaller and Jürgen Wolfart, *Semi-arithmetic Fuchsian groups and modular embeddings*, J. London Math. Soc. (2) **61** (2000), no. 1, pp. 13–24.
- [43] Mika Seppälä and Tuomas Sorvali, *Traces of commutators of Möbius transformations*, Math. Scand. **68** (1991), no. 1, pp. 53–58.
- [44] Jean-Pierre Serre, *Géométrie algébrique et géométrie analytique*, Ann. Inst. Fourier, Grenoble **6** (1955–1956), pp. 1–42.
- [45] ———, *Cohomologie galoisienne*, Cours au Collège de France, vol. 1962, Springer-Verlag, Berlin-Heidelberg-New York, 1962/1963.
- [46] Kuang-yen Shih, *Conjugations of arithmetic automorphic function fields*, Invent. Math. **44** (1978), no. 1, pp. 87–102.
- [47] Goro Shimura, *Construction of class fields and zeta functions of algebraic curves*, Ann. of Math. (2) **85** (1967), pp. 58–159.
- [48] Carlos T. Simpson, *Constructing variations of Hodge structure using Yang-Mills theory and applications to uniformization*, J. Amer. Math. Soc. **1** (1988), no. 4, pp. 867–918.
- [49] ———, *Harmonic bundles on noncompact curves*, J. Amer. Math. Soc. **3** (1990), no. 3, pp. 713–770.
- [50] ———, *Higgs bundles and local systems*, Inst. Hautes Études Sci. Publ. Math. (1992), no. 75, pp. 5–95.
- [51] Kisao Takeuchi, *Arithmetic triangle groups*, J. Math. Soc. Japan **29** (1977), no. 1, pp. 91–106.
- [52] Gerard van der Geer, *Hilbert modular surfaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 16, Springer-Verlag, Berlin, 1988.
- [53] Eckart Viehweg and Kang Zuo, *A characterization of certain Shimura curves in the moduli stack of abelian varieties*, J. Differential Geom. **66** (2004), no. 2, pp. 233–287.

UNIVERSITÄT BONN, MATHEMATISCHES INSTITUT, ENDENICHER ALLEE 60, 53115 BONN, GERMANY

*E-mail address:* robert.a.kucharczyk@googlemail.com